

A2

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
29 March 2001 (29.03.2001)

PCT

(10) International Publication Number  
**WO 01/21772 A2**

- (51) International Patent Classification<sup>7</sup>: C12N 9/00
- (21) International Application Number: PCT/US00/25993
- (22) International Filing Date:  
21 September 2000 (21.09.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
09/400,494 21 September 1999 (21.09.1999) US  
60/210,072 7 June 2000 (07.06.2000) US  
60/221,836 28 July 2000 (28.07.2000) US  
60/227,860 24 August 2000 (24.08.2000) US
- (71) Applicant: OMNIGENE BIOPRODUCTS [US/US];  
763(d) Concord Avenue, Cambridge, MA 02138 (US).
- (72) Inventors: YOCUM, Rogers, R.; 4 Orchard Lane, Lexington, MA 02420 (US). PATTERSON, Thomas, A.; 89 Church Street, North Attleboro, MA 02760 (US); HERMANN, Theron; 18 Chilhowie Drive, Kinnelon, NJ 07405 (US). PERO, Janice, G.; 20 Solomon Pierce Road, Lexington, MA 02420 (US).
- (74) Agents: HANLEY, Elizabeth, A. et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).
- (81) Designated States (*national*): AU, BR, CA, CN, CZ, HU, ID, IL, IN, JP, KR, LT, LV, MX, NO, PL, RU, SG, SK, TR, UA, ZA.
- (84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
- Published:  
— Without international search report and to be republished upon receipt of that report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: METHODS AND MICROORGANISMS FOR PRODUCTION OF PANTO-COMPOUNDS

(57) Abstract: The present invention features methods of producing panto-compounds (e.g., pantothenate) using microorganisms in which the pantothenate biosynthetic pathway and/or the isoleucine-valine biosynthetic pathway and/or the coenzymeA biosynthetic pathway has been manipulated. Methods featuring ketopantoate reductase overexpressing microorganisms as well as aspartate  $\alpha$ -decarboxylase overexpressing microorganisms are provided. Methods of producing panto-compounds in a precursor-independent manner and in high yield are described. Recombinant microorganisms, vectors, isolated nucleic acid molecules, genes and gene products useful in practicing the above methodologies are also provided. The present invention also features a previously unidentified microbial pantothenate kinase gene, *coaX*, as well as methods of producing panto-compounds utilizing microorganisms having modified pantothenate kinase activity. Recombinant microorganisms, vectors, isolated *coaX* nucleic acid molecules and purified CoaX proteins are featured. Also featured are methods for identifying pantothenate kinase modulators utilizing the recombinant microorganisms and/or purified CoaX proteins of the present invention.

## METHODS AND MICROORGANISMS FOR PRODUCTION OF PANTO-COMPOUNDS

### Background of the Invention

5 Pantothenate, also known as pantothenic acid or vitamin B5, is a member of the B complex of vitamins and is a nutritional requirement for mammals, including livestock and humans (*e.g.*, from food sources, as a water soluble vitamin supplement or as a feed additive). In cells, pantothenate is used primarily for the biosynthesis of coenzyme A (CoA) and acyl carrier protein (ACP). These coenzymes function in the  
10 metabolism of acyl moieties which form thioesters with the sulfhydryl group of the 4'-phosphopantetheine portion of these molecules. These coenzymes are essential in all cells, participating in over 100 different intermediary reactions in cellular metabolism.

The conventional means of synthesizing pantothenate (in particular, the bioactive D isomer) is *via* chemical synthesis from bulk chemicals, a process which is hampered  
15 by excessive substrate cost as well as the requirement for optical resolution of racemic intermediates (*e.g.*, resolution of DL-pantolactone to obtain D-pantolactone for chemical condensation with  $\beta$ -alanine). Accordingly, researchers have recently looked to bacterial or microbial systems that produce enzymes useful in pantothenate biosynthesis processes (as bacteria are themselves capable of synthesizing pantothenate). In  
20 particular, bioconversion processes have been evaluated as a means of favoring production of the D isomer of pantothenic acid, *e.g.*, using microorganisms which selectively hydrolyze a DL-pantothenic acid ester to D-pantothenic acid; microorganisms which selectively decompose L-pantolactone resulting in D-pantolactone alone; and microorganisms which selectively hydrolyze DL-pantolactone  
25 to D-pantoic acid.

There is still, however, significant need for improved pantothenate production processes, in particular, for processes requiring reduced quantities of substrates and/or less expensive substrates. To this end, methods of direct microbial synthesis have recently been examined as a means of improving D-pantothenate production. In  
30 microbes, pantothenate biosynthesis is a multistep pathway resulting in condensation of pantoate (derived from  $\alpha$ -ketoisovalerate) and  $\beta$ -alanine to form D-pantothenate. The isoleucine-valine (*ilv*) pathway biosynthetic enzymes, acetohydroxyacid synthetase (the *ilvBN* or *alsS* gene product), acetohydroxyacid isomeroreductase (the *ilvC* gene product) and dihydroxyacid dehydratase (the *ilvD* gene product) catalyze the conversion of  
35 pyruvate to  $\alpha$ -ketoisovalerate. The reactions are further catalyzed by the pantothenate (*pan*) pathway biosynthetic enzymes ketopantoate hydroxymethyltransferase (the *panB* gene product), ketopantoate reductase (the *panE* gene product), aspartate- $\alpha$ -

decarboxylase (the *panD* gene product) and pantothenate synthetase (the *panC* gene product).

The genes encoding the enzymes involved in the biosynthesis of pantothenic acid in *Salmonella typhimurium* and *Escherichia coli* have recently been identified and characterized (Frodyma and Downs (1998) *J. Biol. Chem.* 273:5572-5576 and Jackowski (1996) pp. 687-694, In Neidhardt *et al* (ed.) *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, 2<sup>nd</sup> ed. *Am. Soc. Microbiol.* Wash, D.C). In *E. coli*, for example, the biosynthesis of pantothenic acid consists of four key steps. The first reaction is catalyzed by the *panB* gene product, ketopantoate hydroxymethyltransferase, and uses the L-valine intermediate  $\alpha$ -ketoisovalerate to generate ketopantoate, which is subsequently reduced to pantoate by the *panE* gene product, ketopantoate reductase. The *panD* gene product, aspartate- $\alpha$ -decarboxylase, generates  $\beta$ -alanine from aspartate. The *panC* gene product, pantothenate synthetase, subsequently ligates  $\beta$ -alanine with pantoate to yield D-pantothenate.

The authors Dusch *et al.* described the identification of the *Corynebacterium glutamicum panD* gene and reported that expression of the *C. glutamicum panD* gene in *E. coli* yielded a strain producing pantothenate with a specific productivity of 140 ng of pantothenate per mg (dry weight) per hour. (Dusch *et al.* (1999) *Appl. Environ. Microbiol.* 65:1530-1539).

The authors Sahm and Eggeling have further identified the *Corynebacterium glutamicum panB* and *panC* genes and have described a genetically engineered strain of *C. glutamicum* which overexpresses the *panBC* genes (Sahm and Eggeling (1999) *Appl. Environ. Microbiol.* 65:1973-1979). The engineered strain produces pantothenate, however, it was necessary to overexpress the genes responsible for  $\alpha$ -ketoisovalerate production in the host organism in order that pantothenic acid production could be detected. Moreover, without the addition of  $\beta$ -alanine, no substantial amounts of pantothenate accumulated with the strain constructed.

Likewise, a method of producing D-pantothenic acid has been described that takes advantage of a sodium salicylate resistant mutant strain of *E. coli* which produces D-pantothenic acid when cultured in the presence of  $\beta$ -alanine (U.S. Patent No. 5,518,906). Generation of *E. coli* strains resistant to  $\alpha$ -ketoisovaleric acid and/or  $\alpha$ -ketobutyric acid, and/or  $\alpha$ -aminobutyric acid, and/or  $\beta$ -hydroxyaspartic acid and/or O-methyl-threonine, in addition to salicylic acid, further increased pantothenic acid production. Moreover, transformation of a plasmid DNA carrying the *panB*, *panC* and *panD* genes into the salicylic acid resistant mutant strain resulted in increased pantothenate production, however, up to 20 g/L  $\beta$ -alanine or more was fed in the examples given. The *panB-panC-panD* genes are clustered on the *E. coli* chromosome.

Finally, a method of producing D-pantothenic acid has been described which utilizes a salicylic acid-resistant,  $\alpha$ -ketoisovalerate-resistant,  $\alpha$ -ketobutyrate-resistant,  $\beta$ -hydroxyaspartate-resistant, o-methylthreonine-resistant *E. coli* strain transformed with pantothenate biosynthesis gene-containing DNA fragments and/or branched amino acid biosynthesis gene-containing DNA fragments and cultured in the presence of  $\beta$ -alanine (U.S. Patent No. 5,932,457).

Pantothenate production in bacteria results from the condensation of pantoate and  $\beta$ -alanine and involves the pantothenate biosynthetic enzymes ketopantoate hydroxymethyltransferase (the *panB* gene product), ketopantoate reductase (the *panE* gene product), aspartate- $\alpha$ -decarboxylase (the *panD* gene product) and pantothenate synthetase (the *panC* gene product). Although pantothenate is biologically active as a vitamin, it is further metabolized in all cells to Coenzyme A (CoA) which participates as an acyl group carrier in the tricarboxylic acid (TCA) cycle, fatty acid metabolism and numerous other reactions of intermediary metabolism. The initial (and possibly rate-controlling) step in the conversion of pantothenate to Coenzyme A (CoA) is phosphorylation of pantothenate by pantothenate kinase. A pantothenate kinase activity was first identified in *Salmonella typhimurium* by screening for temperature-sensitive mutants which synthesized CoA at permissive temperatures but excreted pantothenate at non-permissive temperatures. The mutations were mapped in the *Salmonella* chromosome and the genetic locus was designated *coaA*. The gene encodes the enzyme that catalyzes the first step in the biosynthesis of coenzyme A from pantothenate (Dunn and Snell (1979) *J. Bacteriol.* 140:805-808). *Escherichia coli* temperature sensitive mutants have also been isolated and characterized (Vallari and Rock (1987) *J. Bacteriol.* 169:5795-5800). These mutants (named *coaA15(Ts)*) are defective in the conversion of pantothenate to CoA and further exhibit a temperature-sensitive growth phenotype, indicating that pantothenate kinase activity is essential for growth. Moreover, it was noted that CoA inhibited pantothenate kinase activity to the same degree in the mutant as compared to the wild-type enzyme.

Feedback resistant *E. coli* mutants (named *coaA16(Fr)*) have also been isolated that possess a pantothenate kinase activity that is refractory to feedback inhibition by CoA (Vallari and Jackowski (1988) *J. Bacteriol.* 170:3961-3966). The mutation responsible for the reversion is, surprisingly, not genetically linked to the *coaA* gene by transduction. Additional data described therein support the view that the total cellular CoA content is controlled by both modulation of biosynthesis at the pantothenate kinase step and possibly by degradation of CoA to 4'-phosphopantetheine.



The wild-type *E. coli coaA* gene was cloned by functional complementation of *E. coli* temperature-sensitive mutants. The sequence of the wild-type gene was determined (Song and Jackowski (1992) *J. Bacteriol.* 174:6411-6417 and Flamm *et al.* (1988) *Gene (Amst.)* 74:555-558). Strains containing multiple copies of the *coaA* gene possessed 76-fold higher specific activity of pantothenate kinase, however, there was only a 2.7-fold increase in the steady state level of CoA (Song and Jackowski, *supra*). It has further been reported that the prokaryotic enzyme (encoded by *coaA* in *E. coli* and a variety of other microorganisms) is feedback inhibited by CoA both *in vivo* and *in vitro* with CoA being about five times more potent than acetyl-CoA in inhibiting the enzyme (Song and Jackowski, *supra* and Vallari *et al.*, *supra*). Moreover, it has been reported that the *panB* gene product in *E. coli* is inhibited by CoA (Powers and Snell (1976) *J. Biol. Chem.* 251:3786-3793). These data further support the view that feedback inhibition of pantothenate kinase activity is a critical factor controlling intracellular CoA concentration.

Using standard search and alignment tools, *coaA* homologues have been identified in *Hemophilus influenzae*, *Mycobacterium tuberculosis*, *Vibrio cholerae*, *Streptococcus pyogenes* and *Bacillus subtilis*. By contrast, proteins with significant similarity could not be identified in eukaryotic cells including *Saccharomyces cerevisiae* or in mammalian expressed sequence tag (EST) databases. Using a genetic selection strategy, a cDNA encoding pantothenate kinase activity has recently been identified from *Aspergillus nidulans* (Calder *et al.* (1999) *J. Biol. Chem.* 274:2014-2020). The eukaryotic pantothenate kinase gene (*panK*) has distinct primary structure and unique regulatory properties that clearly distinguish it from its prokaryotic counterpart. A mammalian pantothenate kinase gene (*mpanK1a*) has also been isolated which encodes a protein having homology to the *A. nidulans* PanK protein and to the predicted gene product of GenBank™ Accession Number 927798 identified in the *S. cerevisiae* genome (Rock *et al.* (2000) *J. Biol. Chem.* 275:1377-1383).

### Summary of the Invention

The present invention is based, at least in part, on the discovery of key enzyme-encoding genes of the pantothenate biosynthetic pathway in *Bacillus subtilis*. In particular, the present inventors have identified the *panE* gene of *B. subtilis*. Overexpression or deregulation of the *panE* gene in *B. subtilis* results in enhanced production of the *panE* gene product, ketopantoate reductase, further resulting in increased production of pantothenate. Likewise, mutations in this gene reduce pantothenate production in *B. subtilis* >90%. The present inventors have further identified the presumptive *panBCD* operon in *B. subtilis*, overexpression or

deregulation of which results in increased pantothenate production. The present inventors have further demonstrated that overexpression or deregulation of the *panD* gene in *B. subtilis* (resulting in enhanced production of the *panD* gene product, aspartate- $\alpha$ -decarboxylase) further results in increased production of pantothenate, in particular, in combination with deregulation of genes encoding key enzymes of the isoleucine-valine (*ilv*) biosynthetic pathway.

Accordingly, the present invention features methods of producing pantothenate, as well as other compounds of the pantothenate biosynthetic pathway (*e.g.*, ketopantoate, pantoate and  $\beta$ -alanine), termed "panto-compounds" herein, using microorganisms in which the pantothenate biosynthetic pathway and/or isoleucine-valine biosynthetic pathway has been manipulated such that pantothenate or other desired panto-compounds are produced. In one embodiment, the invention features a method of producing a panto-compound (*e.g.*, pantothenate or pantoate) that involves culturing a microorganism which overexpresses the *panE* gene product, ketopantoate reductase, also referred to herein as a ketopantoate reductase-overexpressing or "KPAR-O" microorganism, under conditions such that the panto-compound (*e.g.*, pantothenate or pantoate) is produced. In another embodiment, the present invention features a method of producing panto-compounds (*e.g.*, pantothenate or pantoate) which includes culturing a microorganism which overexpresses at least one pantothenate biosynthetic enzyme (*e.g.*, at least one of the *panB*, *panC* or *panD* gene products), preferably in a KPAR-O microorganism, under conditions such that the panto-compound (*e.g.*, pantothenate or pantoate) is produced.

Yet another aspect of the invention features methods of producing panto-compounds which are independent of the need to feed precursors (*e.g.*,  $\beta$ -alanine or aspartate and/or  $\alpha$ -ketoisovalerate or valine). In one embodiment, the invention features a method of producing pantothenate in a manner independent of precursor feed that includes culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism having a deregulated isoleucine-valine (*ilv*) pathway under conditions such that pantothenate is produced. In another embodiment, the invention features a method of producing pantothenate in a manner independent of precursor feed that includes culturing an A $\alpha$ D-O microorganism having a deregulated pantothenate (*pan*) pathway and a deregulated isoleucine-valine (*ilv*) pathway, under conditions such that pantothenate is produced. In another embodiment, the invention features a method of producing pantothenate in a manner independent of aspartate or  $\beta$ -alanine feed that includes culturing an A $\alpha$ D-O microorganism under conditions such that pantothenate is produced. In another embodiment, the invention features a method of producing pantothenate in a manner independent of valine or  $\alpha$ -ketoisovalerate feed that includes

culturing a microorganism having a deregulated isoleucine-valine (*ilv*) biosynthetic pathway under conditions such that pantothenate is produced. In yet another embodiment, the invention features a high yield production method for producing pantothenate that includes culturing a manipulated microorganism under conditions  
5 such that pantothenate is produced at a significantly high yield (e.g., at a level greater than 10 g/L, 20 g/L, 30 g/L or 40g/L).

The methods of the present invention further feature microorganisms that overexpresses acetohydroxyacid synthetase or acetohydroxyacid isomeroreductase (e.g., microorganisms transformed with a vector that includes an *ilvBNC* nucleic acid  
10 sequence), microorganisms that overexpresses dihydroxyacid dehydratase (e.g., microorganisms transformed with a vector that includes an *ilvD* nucleic acid sequence), microorganisms that overexpresses aspartate- $\alpha$ -decarboxylase (e.g., microorganisms transformed with a vector that includes a *panD* nucleic acid sequence), microorganisms having a deregulated isoleucine-valine (*ilv*) biosynthetic pathway and microorganisms  
15 having a deregulated pantothenate biosynthetic pathway (e.g., microorganisms that overexpress any of ketopantoate hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase and aspartate- $\alpha$ -decarboxylase, for example, microorganisms transformed with a vector comprising a *panBCD* nucleic acid sequence or a vector comprising a *panEI* nucleic acid sequence). In one embodiment, the recombinant  
20 microorganism is Gram positive (e.g., microorganisms belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*). In another embodiment, the recombinant microorganism is Gram negative. Particularly preferred is a *Bacillus* recombinant microorganism (e.g., *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus subtilis*, *Bacillus pumilus*, *Bacillus halodurans*, and the like). Recombinant  
25 vectors that contain the genes encoding *Bacillus* pantothenate and/or isoleucine-valine biosynthetic enzymes (e.g., *B. subtilis* pantothenate and/or isoleucine-valine biosynthetic enzymes) are also described.

Also featured are methods of producing  $\beta$ -alanine that include culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (AaD-O) microorganism under conditions  
30 such that  $\beta$ -alanine is produced and methods of producing  $\beta$ -alanine that involve contacting a composition comprising aspartate with an isolated *Bacillus* aspartate- $\alpha$ -decarboxylase enzyme under conditions such that  $\beta$ -alanine is produced.

The production methods of the present invention further can include recovering the panto-compound (e.g., pantothenate or pantoate).

35 The present invention further features recombinant microorganisms (e.g., AaD-O microorganisms, microorganisms having a deregulated isoleucine-valine (*ilv*) pathway, microorganisms overexpressing at least one of ketopantoate

hydroxymethyltransferase (the *panB* gene product), pantothenate synthetase (the *panC* gene product), aspartate- $\alpha$ -decarboxylase (the *panD* gene product), ketopantoate reductase (the *panE1* gene product) and microorganisms having a deregulated *panBCD* operon. Also featured are *panB*, *panC*, *panD*, *panE*, *ilvB*, *ilvN*, *alsS*, *ilvC*, and/or *ilvD* nucleic acid molecules, as well as vectors including such nucleic acid molecules and gene products encoded by such nucleic acid molecules.

The methodology of the present invention further includes, for example in addition to overexpressing at least one pantothenate biosynthetic enzyme, deleting or mutating a second pantothenate biosynthetic enzyme, said second pantothenate biosynthetic enzyme preferably being downstream of the desired product in the pantothenate biosynthetic pathway. For example, mutating *panC*, in addition to overexpressing the *panE* gene product, results in even further enhanced or increased production of pantoate. Accordingly, in one embodiment, the invention features a method of producing pantoate which includes culturing a microorganism which overexpresses the *panE* gene product and which has a deletion in the *panC* gene. In another embodiment, the invention features a method of producing pantoate which includes culturing a microorganism which overexpresses the *panE* gene product and/or *panB* gene product and which has a deletion in the *panC* gene. Other exemplary embodiments include a method of producing ketopantoate which includes culturing a microorganism which overexpresses the *panB* gene product and which has a deletion in the *panE* gene and a method of producing  $\beta$ -alanine which includes culturing a microorganism which overexpresses the *panD* gene product and which has a deletion in the *panC* gene. Also included are methods of producing panto-compounds which include overexpressing at least one valine biosynthetic enzyme in a microorganism which has at least one pantothenate biosynthetic enzyme deleted.

The present invention is also based at least in part, on the identification and characterization of a previously unidentified microbial pantothenate kinase gene, *coaX*. *CoaX* was first identified in *Bacillus subtilis* and corresponds to an open reading frame in a portion of the chromosomal DNA that includes the 5' end of the *ftsH* gene, and all of the *yacB*, *yacC*, *yacD*, *cysK* and *pabB* genes. The present inventors have demonstrated that the *yacB* open reading frame encodes a novel pantothenate kinase activity, the gene being unrelated by homology to any previously known pantothenate kinase gene. The gene has been renamed *coaX*, as it encodes the enzyme which catalyzes the first step in the pathway from pantothenate to CoaA.

Accordingly, the present invention features new and improved methods of producing pantothenate and other key compounds of the pantothenate biosynthetic pathway (e.g., panto-compounds) utilizing microorganisms having modified

pantothenate kinase activity. In particular, the present invention features recombinant microorganisms that contain the *coaX* gene or that contain a mutant *coaX* gene, having reduced pantothenate kinase activity. In one embodiment, the invention features such recombinant microorganisms further having a deregulated pantothenate biosynthetic pathway. In another embodiment, the invention features such recombinant microorganisms further having a deregulated isoleucine-valine (*ilv*) pathway. In a preferred embodiment, the microorganisms belong to the genus *Bacillus* (e.g., *B. subtilis*).

The present invention also features recombinant microorganisms (e.g., microorganisms belonging to the genus *Bacillus*, for example, *B. subtilis*) that contain the *coaA* gene or that contain a mutant *coaA* gene, optionally including a *coaX* gene or mutant thereof, having reduced pantothenate kinase activity. In one embodiment, the invention features such recombinant microorganisms further having a deregulated pantothenate biosynthetic pathway or having a deregulated isoleucine-valine (*ilv*) pathway.

Also featured are vectors that contain isolated *coaX* or *coaA* genes as well as mutant *coaX* and/or *coaA* genes. Isolated nucleic acid molecules that contain isolated *coaX* genes or mutant *coaX* genes are featured in addition to isolated CoaX proteins and mutant CoaX proteins.

The nucleic acids, vectors and recombinant microorganisms described above are particularly useful in the methodologies of the present invention. In particular, the invention features methods of enhancing panto-compound production (e.g., ketopantoate, pantoate and or pantothenate production) that include culturing a recombinant microorganism having a mutant *coaX* gene under conditions such that panto-compound production is enhanced. In one embodiment, the recombinant microorganism further includes a mutant *coaA* gene. In another embodiment, the recombinant microorganism further includes a mutant *avtA* and/or mutant *ilvE* gene and/or mutant *ansB* gene and/or mutant *alsD* gene. Also featured are methods for identifying pantothenate modulators utilizing the recombinant microorganisms and purified CoaX proteins of the present invention.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

### **Brief Description of the Drawings**

*Figure 1* is a schematic representation of the pantothenate biosynthetic pathway.

*Figure 2* is a schematic representation of the plasmid pAN240, containing sequences ligated upstream of the  $P_{26}$ *panBCD* cassette, equivalent to the integrated  
5 version in strain PA221.

*Figure 3A* is a schematic representation of the plasmid pAN004, containing the *panBCD* operon expressed from  $P_{26}$  and RBS1.

*Figure 3B* is a schematic representation of the plasmid pAN006, containing the *panBCD* operon expressed from  $P_{26}$  and RBS2.

10 *Figure 4* is a schematic representation of the plasmid pAN236, containing an integratable and amplifiable  $P_{26}$ -RBS2-*panE1* expression cassette.

*Figure 5* is a schematic representation of the construction of plasmid pAN423.

*Figure 6* is a schematic representation of the construction of plasmids pAN426 and pAN427.

15 *Figure 7* is a schematic representation of the construction of plasmids pAN428 and pAN429.

*Figure 8* is a schematic representation of the construction of plasmid pAN431.

*Figure 9* is a schematic representation of the construction of plasmid pAN441.

*Figure 10* is a schematic representation of the construction of plasmid pAN440.

20 *Figure 11* is a schematic representation of the plasmid pAN251 designed to integrate a single copy of a  $P_{26}$ -*panE1* cassette at the *panE1* locus by double crossover.

*Figure 12* is a schematic representation of the plasmid pAN267 designed to integrate a single copy of a  $P_{26}$ -*ilvBNC* cassette at the *amyE* locus.

25 *Figure 13* is a schematic representation of the plasmid pAN257, a clone of *Bacillus subtilis ilvD* in a low copy vector.

*Figure 14* is a schematic representation of the plasmid pAN263, designed to integrate a single copy of a  $P_{26}$ -*ilvD* cassette at the *ilvD* locus.

*Figure 15* is a schematic representation of the plasmid pAN261, designed to disrupt the *Bacillus subtilis ilvD* gene with the *cat* gene.

30 *Figure 16* is a schematic representation of the Coenzyme A biosynthetic pathway in *E. coli*.

*Figure 17* is a schematic representation of the structure of pAN296, a plasmid designed to delete most of the *B. subtilis coaA* gene and substitute a chloramphenicol resistance gene.

35 *Figure 18* is a schematic representation of the structure of the *Bacillus subtilis* genome in the region of the *coaA* gene. The scale is in base pairs and the significant open reading frames are shown by open arrows.

Figure 19 is a schematic representation of the plasmid pAN281, a plasmid for expressing *Bacillus subtilis* *coaA* after integration at the *bpr* locus.

Figure 20A-B depicts a multiple sequence alignment (MSA) of the amino acid sequences encoded by six known or predicted microbial *coaA* genes. SEQ ID NOs:4-6 and 1-3 correspond to the amino acid sequences of *Mycobacterium leprae* (SwissProt™ Accession No. Q9X795), *Mycobacterium tuberculosis* (SwissProt™ Accession No. O53440), *Streptomyces coelicolor* (SwissProt™ Accession No. O86799), *Haemophilus influenzae* (SwissProt™ Accession No. P44793), *Escherichia coli* SwissProt™ Accession No. P15044) and *Bacillus subtilis* (SwissProt™ Accession No. P54556), respectively. The alignment was generated using ClustalW MSA software at the GenomeNet CLUSTALW Server at the Institute for Chemical Research, Kyoto University. The following parameters were used: Pairwise Alignment, K-tuple (word) size = 1, Window size = 5, Gap Penalty = 3, Number of Top Diagonals = 5, Scoring Method = Percent; Multiple Alignment, Gap Open Penalty = 10, Gap Extension Penalty = 0.0, Weight Transition = No, Hydrophilic residues = Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg and Lys, Hydrophobic Gaps = Yes; and Scoring Matrix = BLOSUM.

Figure 21 is a schematic representation of the structure of the *Bacillus subtilis* genome in the region of the *coaX* (*yacB*) gene. The scale is in base pairs, the significant open reading frames are shown by open arrows and certain predicted restriction fragments are indicated by thick bars.

Figure 22 is a schematic representation of the structure of pAN341 and pAN342, two independent PCR-derived clones of *B. subtilis* *yacB* (remained herein as *coaX*).

Figure 23A-D depicts a multiple sequence alignment (MSA) of the amino acid sequences encoded by fourteen known or predicted microbial *coaX* genes. SEQ ID NOs:9, 74, 7-8, 75, 11, 10 and 12-18 correspond to the amino acid sequences of *Bacillus subtilis* (SwissProt™ Accession No. P37564), *Clostridium acetobutylicum* (WIT™ Accession No. RCA03301, Argonne National Laboratories), *Streptomyces coelicolor* (PIR™ Accession No. T36391), *Mycobacterium tuberculosis* (SwissProt™ Accession No. O06282), *Rhodobacter capsulatus* (WIT™ Accession No. RRC02473), *Desulfovibrio vulgaris* (DBJ™ Accession No. BAA21476.1), *Deinococcus radiodurans* (SwissProt™ Accession No. Q9RX54), *Thermotoga maritima* (GenBank™ Accession No. AAD35964.1), *Treponema pallidum* (SwissProt™ Accession No. O83446), *Borrelia burgdorferi* (SwissProt™ Accession No. O51477), *Aquifex aeolicus* (SwissProt™ Accession No. O67753), *Synechocystis* sp. (SwissProt™ Accession No. P74045), *Helicobacter pylori* (SwissProt™ Accession No. O25533), and *Bordetella pertussis* (SwissProt™ Accession No. Q45338), respectively. The alignment was generated using ClustalW MSA software at the GenomeNet CLUSTALW Server at the

Institute for Chemical Research, Kyoto University. The following parameters were used: Pairwise Alignment, K-tuple (word) size = 1, Window size = 5, Gap Penalty = 3, Number of Top Diagonals = 5, Scoring Method = Percent; Multiple Alignment, Gap Open Penalty = 10, Gap Extension Penalty = 0.0, Weight Transition = No, Hydrophilic residues = Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg and Lys, Hydrophobic Gaps = Yes; and Scoring Matrix = BLOSUM.

Figure 24 depicts a multiple sequence alignment of a portion of the protein sequences of the *coaA* gene products from the following microorganisms: *Bacillus subtilis*, *Escherichia coli*, *Haemophilus influenzae*, *Mycobacterium leprae*, *Mycobacterium tuberculosis*, and *Streptomyces coelicolor*. The residues that are mutated in *E. coli coaA15(Ts)* and *B. subtilis coaA282A* are indicated below and above the alignment, respectively. The portions correspond to amino acid residues 168-187 of SEQ ID NO:3, 167-186 of SEQ ID NO:2, 165-184 of SEQ ID NO:1, 169-188 of SEQ ID NO:4, 169-188 of SEQ ID NO:5 and 179-198 of SEQ ID NO:6, respectively.

Figure 25 is a schematic representation of the structure of pAN294, a plasmid for integrating mutagenized *B. subtilis coaA* at its native locus.

Figure 26 is a schematic representation of the structure of pAN336, a plasmid designed to delete *B. subtilis coaX* from its chromosomal locus and replace it with a kanamycin resistance gene.

### **Detailed Description of the Invention**

The present invention features new and improved methods of producing pantothenate and other key compounds of the pantothenate biosynthetic pathway (referred to herein as "panto-compounds", for example, pantothenate, ketopantoate, pantoate and  $\beta$ -alanine) using microorganisms in which the pantothenate biosynthetic pathway has been manipulated such that pantothenate or other desired panto-compounds are produced.

The new and improved methodologies of the present invention include methods of producing panto-compounds (e.g., pantothenate) in microorganisms having at least one enzyme of the pantothenate biosynthetic pathway manipulated such that pantothenate or other desired panto-compounds are produced (e.g., produced at an increased level). For example, the invention features methods of producing panto-compounds (e.g., pantothenate) in microorganisms having at least one of ketopantoate hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase or aspartate- $\alpha$ -decarboxylase manipulated such that pantothenate or other desired panto-compounds are produced. The methodologies of the present invention also include methods of producing panto-compounds (e.g., pantothenate) in microorganisms having at least one



valine-isoleucine biosynthetic enzyme, described herein, manipulated such that pantothenate or other desired panto-compounds are produced. For example, the invention features methods of producing panto-compounds (*e.g.*, pantothenate) in microorganisms having at least one of acetohydroxyacid synthetase, acetohydroxyacid isomeroeductase or dihydroxyacid dehydratase manipulated such that pantothenate or other desired panto-compounds are produced.

The invention also features methods of producing panto-compounds that involve culturing a ketopantoate reductase-overexpressing (KPAR-O) microorganism under conditions such that the panto-compound is produced. The invention also features methods of producing pantothenate in a manner independent of precursor feed that involve culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism under conditions such that pantothenate is produced. Also featured are  $\beta$ -alanine independent high yield pantothenate production methods as well as methods of producing  $\beta$ -alanine. The present invention also features methods for enhancing production of panto-compounds that involve culturing pantothenate kinase mutants. In particular, the present invention features new and improved methods of producing pantothenate and other key compounds of the pantothenate biosynthetic pathway (*e.g.*, panto-compounds) utilizing microorganisms having modified pantothenate kinase activity, for example, microorganisms that include the *coaX* gene or that include a mutant *coaX* gene, having reduced pantothenate kinase activity.

In order that the present invention may be more readily understood, certain terms are first defined herein.

The term "pantothenate biosynthetic pathway" includes the biosynthetic pathway involving pantothenate biosynthetic enzymes (*e.g.*, polypeptides encoded by biosynthetic enzyme-encoding genes), compounds (*e.g.*, precursors, substrates, intermediates or products), cofactors and the like utilized in the formation or synthesis of pantothenate. The term "pantothenate biosynthetic pathway" includes the biosynthetic pathway leading to the synthesis of pantothenate in a microorganisms (*e.g.*, *in vivo*) as well as the biosynthetic pathway leading to the synthesis of pantothenate *in vitro*. Figure 1 includes a schematic representation of the pantothenate biosynthetic pathway. Pantothenate biosynthetic enzymes are depicted in bold and their corresponding genes indicated in italics.

The term "pantothenate biosynthetic enzyme" includes any enzyme utilized in the formation of a compound (*e.g.*, intermediate or product) of the pantothenate biosynthetic pathway. According to Figure 1, synthesis of pantoate from  $\alpha$ -ketoisovalerate ( $\alpha$ -KIV) proceeds *via* the intermediate, ketopantoate. Formation of ketopantoate is catalyzed by the pantothenate biosynthetic enzyme ketopantoate

hydroxymethyltransferase (the *panB* gene product). Formation of pantoate is catalyzed by the pantothenate biosynthetic enzyme ketopantoate reductase (the *panE* gene product). Synthesis of  $\beta$ -alanine from aspartate is catalyzed by the pantothenate biosynthetic enzyme aspartate- $\alpha$ -decarboxylase (the *panD* gene product). Formation of  
5 pantothenate from pantoate and  $\beta$ -alanine (e.g., condensation) is catalyzed by the pantothenate biosynthetic enzyme pantothenate synthetase (the *panC* gene product).

The term "isoleucine-valine biosynthetic pathway" includes the biosynthetic pathway involving isoleucine-valine biosynthetic enzymes (e.g., polypeptides encoded by biosynthetic enzyme-encoding genes), compounds (e.g., precursors, substrates,  
10 intermediates or products), cofactors and the like utilized in the formation or synthesis of conversion of pyruvate to valine or isoleucine. The term "isoleucine-valine biosynthetic pathway" includes the biosynthetic pathway leading to the synthesis of valine or isoleucine in a microorganisms (e.g., *in vivo*) as well as the biosynthetic pathway leading to the synthesis of valine or isoleucine *in vitro*. Figure 1 includes a  
15 schematic representation of the isoleucine-valine biosynthetic pathway. Isoleucine-valine biosynthetic enzymes are depicted in bold italics and their corresponding genes indicated in italics

The term "isoleucine-valine biosynthetic enzyme" includes any enzyme utilized in the formation of a compound (e.g., intermediate or product) of the isoleucine-valine  
20 biosynthetic pathway. According to Figure 1, synthesis of valine from pyruvate proceeds *via* the intermediates, acetolactate,  $\alpha,\beta$ -dihydroxyisovalerate ( $\alpha,\beta$ -DHIV) and  $\alpha$ -ketoisovalerate ( $\alpha$ -KIV). Formation of acetolactate from pyruvate is catalyzed by the isoleucine-valine biosynthetic enzyme acetohydroxyacid synthetase (the *ilvBN* gene product, or alternatively, the *alsS* gene product). Formation of  $\alpha,\beta$ -DHIV from  
25 acetolactate is catalyzed by the isoleucine-valine biosynthetic enzyme acetohydroxyacidisomero reductase (the *ilvC* gene product). Synthesis of  $\alpha$ -KIV from  $\alpha,\beta$ -DHIV is catalyzed by the isoleucine-valine biosynthetic enzyme dihydroxyacid dehydratase (the *ilvD* gene product). Moreover, valine and isoleucine can be interconverted by branched chain amino acid transaminases.

30 As used herein, each of ketopantoate, pantoate,  $\beta$ -alanine and pantothenate are "panto-compounds". The term "panto-compound" includes a compound (e.g., a substrate, intermediate or product) in the pantothenate biosynthetic pathway which is downstream from a particular pantothenate biosynthetic enzyme. In one example, a panto-compound is downstream of the pantothenate biosynthetic enzyme ketopantoate  
35 hydroxymethyltransferase (the *panB* gene product) and can include ketopantoate, pantoate and/or pantothenate. In another example, a panto-compound is downstream of the pantothenate biosynthetic enzyme ketopantoate reductase (the *panE* gene product)

and can include pantoate and/or pantothenate. In yet another example, a panto-compound is downstream of the pantothenate biosynthetic enzyme pantothenate synthetase (the *panC* gene product) and can include pantothenate. In yet another example, a panto-compound is downstream of the pantothenate biosynthetic enzyme aspartate- $\alpha$ -decarboxylase (the *panD* gene product) and can include  $\beta$ -alanine and/or pantothenate.

Preferred panto-compounds include pantothenate and pantoate. The term "pantothenate" includes the free acid form of pantothenate, also referred to as "pantothenic acid" as well as any salt thereof (*e.g.*, derived by replacing the acidic hydrogen of pantothenate or pantothenic acid with a cation, for example, calcium, sodium, potassium, ammonium), also referred to as a "pantothenate salt". The term "panto-compound" also includes alcohol derivatives of pantothenate. Preferred pantothenate salts are calcium pantothenate or sodium pantothenate. A preferred alcohol derivative is pantothenol. Pantothenate salts and/or alcohols of the present invention include salts and/or alcohols prepared *via* conventional methods from the free acids described herein. In another embodiment, calcium pantothenate is synthesized directly by a microorganism of the present invention. A pantothenate salt of the present invention can likewise be converted to a free acid form of pantothenate or pantothenic acid by conventional methodology.

The term "pantoate" includes the free acid form of pantoate, also referred to as "pantoic acid" as well as any salt thereof (*e.g.*, derived by replacing the acidic hydrogen of pantoate or pantoic acid with a cation, for example, calcium, sodium, potassium, ammonium), also referred to as a "pantoate salt". Preferred pantoate salts are calcium pantoate or sodium pantoate. Pantoate salts of the present invention include salts prepared *via* conventional methods from the free acids described herein. A pantoate salt of the present invention can likewise be converted to a free acid form of pantoate or pantoic acid by conventional methodology. Moreover, a free acid form of pantoate or pantoic acid can be converted to pantolactone by conventional methodology.

The term "CoA biosynthetic pathway" includes the biosynthetic pathway involving CoA biosynthetic enzymes (*e.g.*, polypeptides encoded by biosynthetic enzyme-encoding genes), compounds (*e.g.*, precursors, substrates, intermediates or products), cofactors and the like utilized in the formation or synthesis of CoA from pantothenate. A schematic representation of the CoA biosynthetic pathway in *E. coli* is set forth as Figure 16. (The pathway depicted is also presumed to be that utilized by other microorganisms.) The term "CoA biosynthetic pathway" includes the biosynthetic pathway leading to the synthesis of CoA in microorganisms (*e.g.*, *in vivo*) as well as the biosynthetic pathway leading to the synthesis of CoA *in vitro*. The term "Coenzyme A

or CoA biosynthetic enzyme" includes any enzyme utilized in the formation of a compound (e.g., intermediate or product) of the CoA biosynthetic pathway, for example, the *coaA*, *panK* or *coaX* gene product which catalyzes the phosphorylation of pantothenate to form 4'-phosphopantothenate, or the *coaD* gene product which catalyzes the conversion of 4'-phosphopantetheine to dephosphocoenzyme A.

I. Recombinant Microorganisms and Methods for Culturing  
Microorganisms Such That Panto-Compounds are Produced

The methodologies of the present invention feature microorganisms, e.g., recombinant microorganisms, preferably including vectors or genes (e.g., wild-type and/or mutated genes) as described herein and/or cultured in a manner which results in the production of a desired product (e.g. a panto-compound or panto-compounds). The term "recombinant" microorganism includes a microorganism (e.g., bacteria, yeast cell, fungal cell, etc.) which has been genetically altered, modified or engineered (e.g., genetically engineered) such that it exhibits an altered, modified or different genotype and/or phenotype (e.g., when the genetic modification affects coding nucleic acid sequences of the microorganism) as compared to the naturally-occurring microorganism from which it was derived. Preferably, a "recombinant" microorganism of the present invention has been genetically engineered such that it overexpresses at least one bacterial gene or gene product (e.g., a pantothenate or isoleucine-valine biosynthetic enzyme encoding-gene) as described herein, preferably a biosynthetic enzyme encoding-gene included within a recombinant vector as described herein and/or a biosynthetic enzyme expressed from a recombinant vector. The ordinary skilled will appreciate that a microorganism expressing or overexpressing a gene product produces or overproduces the gene product as a result of expression or overexpression of nucleic acid sequences and/or genes encoding the gene product.

The term "manipulated microorganism" includes a microorganism that has been engineered (e.g., genetically engineered) or modified such that the microorganism has at least one enzyme of the pantothenate biosynthetic pathway and/or at least one enzyme of the isoleucine-valine biosynthetic pathway modified such that pantothenate or other desired panto-compounds are produced. Modification or engineering of such microorganisms can be according to any methodology described herein including, but not limited to, deregulation of a biosynthetic pathway and/or overexpression of at least one biosynthetic enzyme. A "manipulated" enzyme (e.g., a "manipulated" biosynthetic enzyme) includes an enzyme, the expression or production of which has been altered or modified such that at least one upstream or downstream precursor, substrate or product

of the enzyme is altered or modified, for example, as compared to a corresponding wild-type or naturally occurring enzyme.

The term "overexpressed" or "overexpression" includes expression of a gene product (*e.g.*, a pantothenate biosynthetic enzyme or isoleucine-valine biosynthetic enzyme) at a level greater than that expressed prior to manipulation of the microorganism or in a comparable microorganism which has not been manipulated. In one embodiment, the microorganism can be genetically manipulated (*e.g.*, genetically engineered) to overexpress a level of gene product greater than that expressed prior to manipulation of the microorganism or in a comparable microorganism which has not been manipulated. Genetic manipulation can include, but is not limited to, altering or modifying regulatory sequences or sites associated with expression of a particular gene (*e.g.*, by adding strong promoters, inducible promoters or multiple promoters or by removing regulatory sequences such that expression is constitutive), modifying the chromosomal location of a particular gene, altering nucleic acid sequences adjacent to a particular gene such as a ribosome binding site or transcription terminator, increasing the copy number of a particular gene, modifying proteins (*e.g.*, regulatory proteins, suppressors, enhancers, transcriptional activators and the like) involved in transcription of a particular gene and/or translation of a particular gene product, or any other conventional means of deregulating expression of a particular gene routine in the art (including but not limited to use of antisense nucleic acid molecules, for example, to block expression of repressor proteins).

In another embodiment, the microorganism can be physically or environmentally manipulated to overexpress a level of gene product greater than that expressed prior to manipulation of the microorganism or in a comparable microorganism which has not been manipulated. For example, a microorganism can be treated with or cultured in the presence of an agent known or suspected to increase transcription of a particular gene and/or translation of a particular gene product such that transcription and/or translation are enhanced or increased. Alternatively, a microorganism can be cultured at a temperature selected to increase transcription of a particular gene and/or translation of a particular gene product such that transcription and/or translation are enhanced or increased.

The term "deregulated" or "deregulation" includes the alteration or modification of at least one gene in a microorganism that encodes an enzyme in a biosynthetic pathway, such that the level or activity of the biosynthetic enzyme in the microorganism is altered or modified. Preferably, at least one gene that encodes an enzyme in a biosynthetic pathway is altered or modified such that the gene product is enhanced or increased. The phrase "deregulated pathway" can also include a biosynthetic pathway in

which more than one gene that encodes an enzyme in a biosynthetic pathway is altered or modified such that the level or activity of more than one biosynthetic enzyme is altered or modified. The ability to "deregulate" a pathway (*e.g.*, to simultaneously deregulate more than one gene in a given biosynthetic pathway) in a microorganism  
5 arises from the particular phenomenon of microorganisms in which more than one enzyme (*e.g.*, two or three biosynthetic enzymes) are encoded by genes occurring adjacent to one another on a contiguous piece of genetic material termed an "operon".

The term "operon" includes a coordinated unit of gene expression that contains a promoter and possibly a regulatory element associated with one or more, preferably at  
10 least two, structural genes (*e.g.*, genes encoding enzymes, for example, biosynthetic enzymes). Expression of the structural genes can be coordinately regulated, for example, by regulatory proteins binding to the regulatory element or by anti-termination of transcription. The structural genes can be transcribed to give a single mRNA that encodes all of the structural proteins. Due to the coordinated regulation of genes  
15 included in an operon, alteration or modification of the single promoter and/or regulatory element can result in alteration or modification of each gene product encoded by the operon. Alteration or modification of the regulatory element can include, but is not limited to removing the endogenous promoter and/or regulatory element(s), adding strong promoters, inducible promoters or multiple promoters or removing regulatory  
20 sequences such that expression of the gene products is modified, modifying the chromosomal location of the operon, altering nucleic acid sequences adjacent to the operon or within the operon such as a ribosome binding site, increasing the copy number of the operon, modifying proteins (*e.g.*, regulatory proteins, suppressors, enhancers, transcriptional activators and the like) involved in transcription of the operon and/or  
25 translation of the gene products of the operon, or any other conventional means of deregulating expression of genes routine in the art (including but not limited to use of antisense nucleic acid molecules, for example, to block expression of repressor proteins). Deregulation can also involve altering the coding region of one or more genes to yield, for example, an enzyme that is feedback resistant or has a higher or lower  
30 specific activity.

A particularly preferred "recombinant" microorganism of the present invention has been genetically engineered to overexpress a bacterially-derived gene or gene product. The term "bacterially-derived" or "derived-from", for example bacteria, includes a gene which is naturally found in bacteria or a gene product (*e.g.*, ketopantoate  
35 hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase, aspartate- $\alpha$ -decarboxylate, acetohydroxyacid synthetase, acetohydroxyacid isomeroreductase or

dihydroxyacid dehydratase) which is encoded by a bacterial gene (e.g., encoded by *panB*, *panE*, *panC*, *panD*, *ilvB*, *ilvN*, *alsS*, *ilvC*, or *ilvD*).

The methodologies of the present invention feature recombinant microorganisms which overexpress at least one of ketopantoate hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase or aspartate- $\alpha$ -decarboxylase. A particularly preferred recombinant microorganism of the present invention has been genetically engineered to overexpress a *Bacillus* (e.g., *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus halodurans*, *Bacillus subtilis*, and *Bacillus pumilus*, etc.) biosynthetic enzyme (e.g., has been engineered to overexpress at least one of *B. subtilis* ketopantoate reductase (the *panE* gene product) (e.g., ketopantoate reductase having the amino acid sequence of SEQ ID NO:30 or encoded by the nucleic acid sequence of SEQ ID NO:29), *B. subtilis* ketopantoate hydroxymethyltransferase (the *panB* gene product) (e.g., ketopantoate hydroxymethyltransferase having the amino acid sequence of SEQ ID NO:24 or encoded by a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:23), *B. subtilis* pantothenate synthetase (the *panC* gene product) (e.g., pantothenate synthetase having the amino acid sequence of SEQ ID NO:26 or encoded by a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:25) and/or *B. subtilis* aspartate- $\alpha$ -decarboxylase (the *panD* gene product) (e.g., aspartate- $\alpha$ -decarboxylase having the amino acid sequence of SEQ ID NO:28 or encoded by a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:27).

In an exemplary embodiment, the invention features a microorganism (e.g., a KPAR-O microorganism) that has been transformed with a vector comprising a *panE* nucleic acid sequence (e.g., a *panE* nucleic acid sequence as set forth in SEQ ID NO:29). In another embodiment, the invention features a microorganism that has been transformed with a vector comprising a *panB* nucleic acid sequence (e.g., a *panB* nucleic acid sequence as set forth in SEQ ID NO:23), a vector comprising a *panC* nucleic acid sequence (e.g., a *panC* nucleic acid sequence as set forth in SEQ ID NO:25) or a vector comprising a *panD* nucleic acid sequence (e.g., a *panD* nucleic acid sequence as set forth in SEQ ID NO:27). In yet another embodiment, the invention features a microorganism having a deregulated *panBCD* operon (e.g., SEQ ID NO:59).

Other preferred "recombinant" microorganisms of the present invention have a deregulated isoleucine-valine (*ilv*) pathway. The phrase "microorganism having a deregulated isoleucine-valine (*ilv*) pathway" includes a microorganism having an alteration or modification in at least one gene encoding an enzyme of the isoleucine-valine (*ilv*) pathway or having an alteration or modification in an operon including more than one gene encoding an enzyme of the isoleucine-valine (*ilv*) pathway. A preferred "microorganism having a deregulated isoleucine-valine (*ilv*) pathway" has been

genetically engineered to overexpress a *Bacillus* (e.g., *B. subtilis*) *ilv* biosynthetic enzyme (e.g., has been engineered to overexpress at least one of acetohydroxyacid synthetase (the *ilvBN* gene products or the *alsS* gene product) (e.g., acetohydroxyacid synthetase having subunits having the amino acid sequences of SEQ ID NO:32 and SEQ ID NO:34 or encoded by nucleic acid molecules having the nucleotide sequence of SEQ ID NO:31 and SEQ ID NO:33 or the nucleotide sequence of SEQ ID NO:58 from nucleotides 1-2246 or acetohydroxyacid synthetase encoded by a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:86), acetohydroxyacid isomeroreductase (the *ilvC* gene product) (e.g., acetohydroxyacid isomeroreductase having the amino acid sequence of SEQ ID NO:36 or encoded by a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:35), dihydroxyacid dehydratase (the *ilvD* gene product) (e.g., dihydroxyacid dehydratase having the amino acid sequence of SEQ ID NO:38 or encoded by a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:37), and/or has been transformed with a vector comprising an *ilvBNC* nucleic acid sequence (SEQ ID NO:58, coding regions from nucleotides 1-1725, 1722-2246 and 2263-3291) and/or an *ilvD* nucleic acid sequence (SEQ ID NO:37).

In another preferred embodiment, a recombinant microorganism is designed or engineered such that a mutant CoaA and/or CoaX biosynthetic enzyme is expressed and at least one pantothenate biosynthetic enzyme and/or at least one isoleucine-valine biosynthetic enzyme is overexpressed or deregulated.

In another preferred embodiment, a microorganism of the present invention overexpresses or is mutated for a gene or biosynthetic enzyme (e.g., a CoA biosynthetic enzyme, pantothenate biosynthetic enzyme or isoleucine-valine biosynthetic enzyme) which is bacterially-derived. The term "bacterially-derived" or "derived-from", for example bacteria, includes a gene product (e.g., ketopantoate hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase, aspartate- $\alpha$ -decarboxylate, acetohydroxyacid synthetase, acetohydroxyacid isomeroreductase, dihydroxyacid dehydratase or pantothenate kinase) which is encoded by a bacterial gene (e.g., *panB*, *panE*, *panC*, *panD*, *ilvBN* (or *alsS*), *ilvC*, *ilvD*, or encoded by *coaA* or *coaX*).

Still other preferred recombinant microorganisms of the present invention are mutant microorganisms. As used herein, the term "mutant microorganism" includes a recombinant microorganism that has been genetically engineered to express a mutated gene or protein that is normally or naturally expressed by the microorganism. Preferably, a mutant microorganism expresses a mutated gene or protein such that the microorganism exhibits an altered, modified or different phenotype (e.g., has been engineered to express a mutated CoaA biosynthetic enzyme, for example, pantothenate kinase). In one embodiment, a mutant microorganism is designed or engineered such



that it includes a mutant *coaX* gene, as defined herein. In another embodiment, a recombinant microorganism is designed or engineered such that it includes a mutant *coaA* gene, as defined herein. In another embodiment, a mutant microorganism is designed or engineered such that a *coaX* gene has been deleted (*i.e.*, the protein encoded by the *coaX* gene is not produced). In another embodiment, a mutant microorganism is designed or engineered such that a *coaA* gene has been deleted (*i.e.*, the protein encoded by the *coaA* gene is not produced). Preferably, a mutant microorganism has a mutant *coaX* gene or a mutant *coaA* gene, or has been engineered to have a *coaX* gene and/or *coaA* deleted, such that that the mutant microorganism encodes a “reduced pantothenate kinase activity”. In the context of a whole microorganism, a “reduced pantothenate kinase activity” can be determined by measuring or assaying for a decrease in an intermediate or product of the CoA biosynthetic pathway, for example, measuring or assaying for 4'-phosphopantothenate, 4'-phosphopantothenylcysteine, 4'-phosphopantetheine, dephosphocoenzyme A, Coenzyme A, apo-acyl carrier protein (apo-ACP) or holo-acyl carrier protein (ACP) in the microorganism (*e.g.*, in a lysate isolated or derived from the microorganism) or in the medium in which the microorganism is cultured (see *e.g.*, Figure 16). Alternatively, a “reduced pantothenate kinase activity” can be determined by measuring or assaying for decreased growth of the microorganism. Alternatively, a “reduced pantothenate kinase activity” can be determined by measuring or assaying for an increase in a panto-compound (*e.g.*, pantothenate) in the microorganism or surrounding media, as panto-compounds lie upstream of the CoA biosynthetic pathway, the first step of which is catalyzed by pantothenate kinase. The invention also features recombinant microorganisms that, in addition to having reduced pantothenate kinase activity (*e.g.*, expressing mutant *coaA* and/or mutant *coaX* genes) have a deregulated pantothenate biosynthesis pathway and/or a deregulated isoleucine-valine (*ilv*) biosynthetic pathway.

In one embodiment, a recombinant microorganism of the present invention is a Gram positive organism (*e.g.*, a microorganism which retains basic dye, for example, crystal violet, due to the presence of a Gram-positive wall surrounding the microorganism). In a preferred embodiment, the recombinant microorganism is a microorganism belonging to a genus selected from the group consisting of *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* and *Streptomyces*. In a more preferred embodiment, the recombinant microorganism is of the genus *Bacillus*. In another preferred embodiment, the recombinant microorganism is selected from the group consisting of *Bacillus subtilis*, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus firmus*, *Bacillus pantothenicus*, *Bacillus amyloliquefaciens*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*,

*Bacillus thuringiensis*, and other Group 1 *Bacillus* species, for example, as characterized by 16S rRNA type (Priest (1993) in *Bacillus subtilis and Other Gram-Positive Bacteria* eds. Sonenshein *et al.*, ASM, Washington, D.C., p. 6). In another preferred embodiment, the recombinant microorganism is *Bacillus brevis* or *Bacillus*  
5 *stearothermophilus*. In another preferred embodiment, the recombinant microorganism is selected from the group consisting of *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus halodurans*, *Bacillus subtilis*, and *Bacillus pumilus*.

In another embodiment, the recombinant microorganism is a Gram negative (excludes basic dye) organism. In a preferred embodiment, the recombinant  
10 microorganism is a microorganism belonging to a genus selected from the group consisting of *Salmonella*, *Escherichia*, *Klebsiella*, *Serratia*, and *Proteus*. In a more preferred embodiment, the recombinant microorganism is of the genus *Escherichia*. In an even more preferred embodiment, the recombinant microorganism is *Escherichia coli*. In another embodiment, the recombinant microorganism is *Saccharomyces* (e.g., *S.*  
15 *cerevisiae*).

An important aspect of the present invention involves culturing the recombinant microorganisms described herein, such that a desired compound (e.g., a desired panto-compound) is produced. The term "culturing" includes maintaining and/or growing a living microorganism of the present invention (e.g., maintaining and/or growing a  
20 culture or strain). In one embodiment, a microorganism of the invention is cultured in liquid media. In another embodiment, a microorganism of the invention is cultured in solid media or semi-solid media. In a preferred embodiment, a microorganism of the invention is cultured in media (e.g., a sterile, liquid media) comprising nutrients essential or beneficial to the maintenance and/or growth of the microorganism (e.g.,  
25 carbon sources or carbon substrate, for example complex carbohydrates such as bean or grain meal, starches, sugars, sugar alcohols, hydrocarbons, oils, fats, fatty acids, organic acids and alcohols; nitrogen sources, for example, vegetable proteins, peptones, peptides and amino acids derived from grains, beans and tubers, proteins, peptides and amino acids derived from animal sources such as meat, milk and animal byproducts such as  
30 peptones, meat extracts and casein hydrolysates; inorganic nitrogen sources such as urea, ammonium sulfate, ammonium chloride, ammonium nitrate and ammonium phosphate; phosphorus sources, for example, phosphoric acid, sodium and potassium salts thereof; trace elements, for example, magnesium, iron, manganese, calcium, copper, zinc, boron, molybdenum, and/or cobalt salts; as well as growth factors such as  
35 amino acids, vitamins, growth promoters and the like).

Preferably, microorganisms of the present invention are cultured under controlled pH. The term "controlled pH" includes any pH which results in production of the desired product (*e.g.*, a panto-compound). In one embodiment, microorganisms are cultured at a pH of about 7. In another embodiment, microorganisms are cultured at a pH of between 6.0 and 8.5. The desired pH may be maintained by any number of methods known to those skilled in the art.

Also preferably, microorganisms of the present invention are cultured under controlled aeration. The term "controlled aeration" includes sufficient aeration (*e.g.*, oxygen) to result in production of the desired product (*e.g.*, panto-compound). In one embodiment, aeration is controlled by regulating oxygen levels in the culture, for example, by regulating the amount of oxygen dissolved in culture media. Preferably, aeration of the culture is controlled by agitating the culture. Agitation may be provided by a propeller or similar mechanical agitation equipment, by revolving or shaking the growth vessel (*e.g.*, fermentor) or by various pumping equipment. Aeration may be further controlled by the passage of sterile air or oxygen through the medium (*e.g.*, through the fermentation mixture). Also preferably, microorganisms of the present invention are cultured without excess foaming (*e.g.*, *via* addition of antifoaming agents).

Moreover, microorganisms of the present invention can be cultured under controlled temperatures. The term "controlled temperature" includes any temperature which results in production of the desired product (*e.g.*, a panto-compound). In one embodiment, controlled temperatures include temperatures between 15°C and 95°C. In another embodiment, controlled temperatures include temperatures between 15°C and 70°C. Preferred temperatures are between 20°C and 55°C, more preferably between 30°C and 45°C or between 30°C and 50°C.

Microorganisms can be cultured (*e.g.*, maintained and/or grown) in liquid media and preferably are cultured, either continuously or intermittently, by conventional culturing methods such as standing culture, test tube culture, shaking culture (*e.g.*, rotary shaking culture, shake flask culture, etc.), aeration spinner culture, or fermentation. In a preferred embodiment, the microorganisms are cultured in shake flasks. In a more preferred embodiment, the microorganisms are cultured in a fermentor (*e.g.*, a fermentation process). Fermentation processes of the present invention include, but are not limited to, batch, fed-batch and continuous methods of fermentation. The phrase "batch process" or "batch fermentation" refers to a closed system in which the composition of media, nutrients, supplemental additives and the like is set at the beginning of the fermentation and not subject to alteration during the fermentation, however, attempts may be made to control such factors as pH and oxygen concentration to prevent excess media acidification and/or microorganism death. The phrase "fed-

batch process" or "fed-batch" fermentation refers to a batch fermentation with the exception that one or more substrates or supplements are added (e.g., added in increments or continuously) as the fermentation progresses. The phrase "continuous process" or "continuous fermentation" refers to a system in which a defined

5 fermentation media is added continuously to a fermentor and an equal amount of used or "conditioned" media is simultaneously removed, preferably for recovery of the desired product (e.g., panto-compound). A variety of such processes have been developed and are well-known in the art.

The phrase "culturing under conditions such that a desired compound (e.g., a  
10 panto-compound, for example, pantothenate) is produced" includes maintaining and/or growing microorganisms under conditions (e.g., temperature, pressure, pH, duration, etc.) appropriate or sufficient to obtain production of the desired compound or to obtain desired yields of the particular compound being produced. For example, culturing is continued for a time sufficient to produce the desired amount of a panto-compound (e.g.,  
15 pantothenate, pantoate or  $\beta$ -alanine). Preferably, culturing is continued for a time sufficient to substantially reach maximal production of the panto-compound. In one embodiment, culturing is continued for about 12 to 24 hours. In another embodiment, culturing is continued for about 24 to 36 hours, 36 to 48 hours, 48 to 72 hours, 72 to 96 hours, 96 to 120 hours, 120 to 144 hours, or greater than 144 hours. In another  
20 embodiment, culturing is continued for a time sufficient to reach production yields of panto-compound, for example, cells are cultured such that at least about 15 to 20 g/L of panto-compound are produced, at least about 20 to 25 g/L panto-compound are produced, at least about 25 to 30 g/L panto-compound are produced, at least about 30 to 35 g/L panto-compound are produced, at least about 35 to 40 g/L panto-compound are  
25 produced (e.g., at least about 37 g/L panto-compound) or at least about 40 to 50 g/L panto compound are produced. In yet another embodiment, microorganisms are cultured under conditions such that a preferred yield of panto-compound, for example, a yield within a range set forth above, is produced in about 24 hours, in about 36 hours, in about 48 hours, in about 72 hours, or in about 96 hours.

30 The methodology of the present invention can further include a step of recovering a desired compound (e.g., a panto-compound). The term "recovering" a desired compound (e.g., a panto-compound) includes extracting, harvesting, isolating or purifying the compound from culture media. Recovering the compound can be performed according to any conventional isolation or purification methodology known  
35 in the art including, but not limited to, treatment with a conventional resin (e.g., anion or cation exchange resin, non-ionic adsorption resin, etc.), treatment with a conventional adsorbent (e.g., activated charcoal, silicic acid, silica gel, cellulose, alumina, etc.),

alteration of pH, solvent extraction (*e.g.*, with a conventional solvent such as an alcohol, ethyl acetate, hexane and the like), dialysis, filtration, concentration, crystallization, recrystallization, pH adjustment, lyophilization and the like. For example, a compound (*e.g.*, a panto-compound) can be recovered from culture media by first removing the microorganisms from the culture. Media is then passed through or over a cation exchange resin to remove unwanted cations and then through or over an anion exchange resin to remove unwanted inorganic anions and organic acids having stronger acidities than the panto-compound of interest (*e.g.*, pantothenate). The resulting panto-compound (*e.g.*, pantothenate) can subsequently be converted to a pantothenate salt (*e.g.*, calcium pantothenate) as described herein.

Preferably, a desired compound of the present invention is "extracted", "isolated" or "purified" such that the resulting preparation is substantially free of other components (*e.g.*, free of media components and/or fermentation byproducts). The language "substantially free of other components" includes preparations of desired compound in which the compound is separated (*e.g.*, purified or partially purified) from media components or fermentation byproducts of the culture from which it is produced. In one embodiment, the preparation has greater than about 80% (by dry weight) of the desired compound (*e.g.*, less than about 20% of other media components or fermentation byproducts), more preferably greater than about 90% of the desired compound (*e.g.*, less than about 10% of other media components or fermentation byproducts), still more preferably greater than about 95% of the desired compound (*e.g.*, less than about 5% of other media components or fermentation byproducts), and most preferably greater than about 98-99% desired compound (*e.g.*, less than about 1-2% other media components or fermentation byproducts). When the desired compound is a panto-compound that has been derivatized to a salt (*e.g.* a pantothenate salt or pantoate salt), the panto-compound is preferably further free (*e.g.*, substantially free) of chemical contaminants associated with the formation of the salt. When the desired compound is a panto-compound that has been derivatized to an alcohol, the panto-compound is preferably further free (*e.g.*, substantially free) of chemical contaminants associated with the formation of the alcohol.

In an alternative embodiment, the desired panto-compound is not purified from the microorganism, for example, when the microorganism is biologically non-hazardous (*e.g.*, safe). For example, the entire culture (or culture supernatant) can be used as a source of product (*e.g.*, crude product). In one embodiment, the culture (or culture supernatant) supernatant is used without modification. In another embodiment, the culture (or culture supernatant) is concentrated. In yet another embodiment, the culture (or culture supernatant) is dried or lyophilized.

## II. Panto-Compound Production Methodologies Featuring Ketopantoate Reductase-Overexpressing Microorganisms

One aspect of the invention features methods of producing a panto-compounds that involve culturing a ketopantoate reductase-overexpressing (KPAR-O) microorganism under conditions such that the panto-compound is produced. The term “ketopantoate reductase-overexpressing (KPAR-O) microorganism” includes a microorganism which has been manipulated such that ketopantoate reductase is overexpressed (e.g., a *B. subtilis* ketopantoate reductase protein having the amino acid sequence of SEQ ID NO:30) and/or has been transformed with a vector comprising a *panE1* nucleic acid sequence (e.g., a *B. subtilis panE1* nucleic acid sequence as set forth in SEQ ID NO:29). In one embodiment, the panto-compound is pantothenate. In another embodiment, the panto-compound is pantoate. In another embodiment, the ketopantoate reductase is bacterial-derived. In another embodiment, the ketopantoate reductase is derived from *Bacillus* (e.g., is derived from *Bacillus subtilis*). In yet another embodiment, the KPAR-O microorganism is Gram positive. In yet another embodiment, the KPAR-O microorganism is a microorganism belonging to a genus selected from the group consisting of *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* and *Streptomyces*. In a preferred embodiment, the KPAR-O microorganism is of the genus *Bacillus*. In a more preferred embodiment, the KPAR-O microorganism is selected from the group consisting of *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus halodurans*, *Bacillus subtilis* and *Bacillus pumilus*. In a particularly preferred embodiment, the KPAR-O microorganism is *Bacillus subtilis*.

In still other embodiments, the KPAR-O microorganism further overexpresses at least one pantothenate biosynthetic enzyme in addition to ketopantoate reductase. In an exemplary embodiment, the KPAR-O microorganism further overexpresses at least one of ketopantoate hydroxymethyltransferase, pantothenate synthetase and aspartate- $\alpha$ -decarboxylase. Also featured are methods of producing panto-compounds, for example, methods that involve culturing a KPAR-O microorganism, which further include the step of recovering the panto-compound.

30

## III. Methods of Producing Panto-Compounds Independent of Precursor Feed Requirements

Depending on the biosynthetic enzyme or combination of biosynthetic enzymes manipulated, it may be desirable or necessary to provide (e.g., feed) microorganisms of the present invention at least one pantothenate biosynthetic precursor such that pantothenate or other desired panto-compounds are produced. The term “pantothenate biosynthetic precursor” or “precursor” includes an agent or compound which, when

35

provided to, brought into contact with, or included in the culture medium of a microorganism, serves to enhance or increase pantothenate biosynthesis. In one embodiment, the pantothenate biosynthetic precursor or precursor is aspartate. In another embodiment, the pantothenate biosynthetic precursor or precursor is  $\beta$ -alanine.

- 5 The amount of aspartate or  $\beta$ -alanine added is preferably an amount that results in a concentration in the culture medium sufficient to enhance productivity of the microorganism (*e.g.*, a concentration sufficient to enhance production of a panto-compound, for example,  $\beta$ -alanine, ketopantoate, pantoate or pantothenate).

- 10 Pantothenate biosynthetic precursors of the present invention can be added in the form of a concentrated solution or suspension (*e.g.*, in a suitable solvent such as water or buffer) or in the form of a solid (*e.g.*, in the form of a powder). Moreover, pantothenate biosynthetic precursors of the present invention can be added as a single aliquot, continuously or intermittently over a given period of time.

- In yet another embodiment, the pantothenate biosynthetic precursor is valine, see *e.g.*, Example III. In yet another embodiment, the pantothenate biosynthetic precursor is  $\alpha$ -ketoisovalerate. Preferably, valine or  $\alpha$ -ketoisovalerate is added in an amount that results in a concentration in the medium sufficient for production of the desired product (*e.g.*, panto-compound) to occur. Pantothenate biosynthetic precursors are also referred to herein as "supplemental pantothenate biosynthetic substrates".

- 20 Providing pantothenate biosynthetic precursors in the pantothenate biosynthetic methodologies of the present invention, can be associated with high costs, for example, when the methodologies are used to produce high yields of panto-compounds. Accordingly, preferred methodologies of the present invention feature microorganisms having at least one biosynthetic enzyme or combination of biosynthetic enzymes (*e.g.*, at least one pantothenate biosynthetic enzyme and/or valine-isoleucine biosynthetic enzyme) manipulated such that pantothenate or other desired panto-compounds are produced in a manner independent of precursor feed. The phrase "a manner independent of precursor feed", for example, when referring to a method for producing a desired compound (*e.g.*, a panto-compound), includes an approach to or a mode of producing the desired compound that does not depend or rely on precursors being provided (*e.g.*, fed) to the microorganism being utilized to produce the desired compound. For example, microorganisms featured in the methodologies of the present invention can be used to produce panto-compounds in a manner requiring no feeding of the precursors aspartate,  $\beta$ -alanine, valine and/or  $\alpha$ -KIV.

- 35 Alternative preferred methodologies of the present invention feature microorganisms having at least one biosynthetic enzyme or combination of biosynthetic enzymes manipulated such that pantothenate or other desired panto-compounds are

produced in a manner substantially independent of precursor feed. The phrase "a manner substantially independent of precursor feed" includes an approach to or a method of producing the desired compound that depends or relies to a lesser extent on precursors being provided (*e.g.*, fed) to the microorganism being utilized. For example, microorganisms featured in the methodologies of the present invention can be used to produce panto-compounds in a manner requiring feeding of substantially reduced amounts of the precursors aspartate,  $\beta$ -alanine, valine and/or  $\alpha$ -KIV. In one embodiment, the invention features methods of producing panto-compounds (*e.g.*, pantothenate) in a manner that requires feeding of less than 5%-10% of the amount of precursor required by a control microorganism (*e.g.*, a microorganism that is dependent, for example is wholly dependent, on precursor feed to efficiently produce the desired compound). In another embodiment, the invention features methods of producing panto-compounds in a manner that requires feeding of less than 15-20% of the amount of precursor required by a control microorganism. In another embodiment, the invention features methods of producing panto-compounds in a manner that requires feeding of less than 25-30%, 35-40%, 45-50% or 55-60% of the amount of precursor required by a control microorganism. As described in Examples I-III herein, particular microorganisms featured in the methodologies of the present invention require, for example, 5 g/L of aspartate,  $\beta$ -alanine, valine or  $\alpha$ -KIV (*e.g.*, in test tube or in shake flask cultures). Accordingly, in a preferred embodiment, the present invention features methods of producing panto-compounds (*e.g.*, pantothenate) in a manner requiring feeding of less than 0.25 g/L, 0.5 g/L, 0.75 g/L, 1 g/L, 1.25 g/L, 1.5 g/L, 1.75 g/L, 2 g/L, 2.25 g/L, 2.5 g/L, 2.75 g/L or 3 g/L.

Preferred methods of producing desired compounds (*e.g.*, panto-compounds) in a manner independent of precursor feed or alternatively, in a manner substantially independent of precursor feed, involve culturing microorganisms which have been manipulated (*e.g.*, designed or engineered, for example, genetically engineered) such that expression of at least one pantothenate biosynthetic enzyme, and/or at least one isoleucine-valine biosynthetic enzyme, is modified. For example, in one embodiment, a microorganism is manipulated (*e.g.*, designed or engineered) such that the production of at least one pantothenate biosynthetic enzyme, and/or at least one isoleucine/valine biosynthetic enzyme is deregulated. In a preferred embodiment, a microorganism is manipulated (*e.g.*, designed or engineered) such that it has a deregulated biosynthetic pathway, for example, a deregulated pantothenate biosynthesis pathway and/or a deregulated isoleucine-valine biosynthetic pathway, as defined herein. In another preferred embodiment, a microorganism is manipulated (*e.g.*, designed or engineered)



such that at least one pantothenate biosynthetic enzyme, and/or at least one isoleucine-valine biosynthetic enzyme is overexpressed.

Preferred methods of producing desired compounds (*e.g.*, panto-compounds) in a manner independent of precursor feed or alternatively, in a manner substantially independent of precursor feed, are as follows. In one embodiment, the invention features a method of producing pantothenate in a manner independent of precursor feed comprising culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism having a deregulated isoleucine-valine (*ilv*) pathway under conditions such that pantothenate is produced. In another embodiment, the invention features a method of producing pantothenate in a manner independent of precursor feed comprising culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism having a deregulated pantothenate (*pan*) pathway and a deregulated isoleucine-valine (*ilv*) pathway, under conditions such that pantothenate is produced. In another embodiment, the invention features a method of producing pantothenate in a manner independent of aspartate or  $\beta$ -alanine feed comprising culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism under conditions such that pantothenate is produced. In yet another embodiment, the invention features a method of producing pantothenate in a manner independent of valine or  $\alpha$ -ketoisovalerate feed comprising culturing a microorganism having a deregulated isoleucine-valine (*ilv*) biosynthetic pathway under conditions such that pantothenate is produced.

The term "aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism" includes a microorganism which has been manipulated such that aspartate- $\alpha$ -decarboxylase is overexpressed. A preferred "aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism" has been transformed with a vector comprising a *B. subtilis* *panD* nucleic acid sequence (*e.g.*, a *panD* nucleic acid sequence that encodes an aspartate- $\alpha$ -decarboxylase protein having the amino acid sequence of SEQ ID NO:28, for example, a *panD* nucleic acid sequence as set forth in SEQ ID NO:27).

The phrase "microorganism having a deregulated isoleucine-valine (*ilv*) pathway" includes a microorganism having an alteration or modification in at least one gene encoding an enzyme of the isoleucine-valine (*ilv*) pathway or having an alteration or modification in an operon including more than one gene encoding an enzyme of the isoleucine-valine (*ilv*) pathway. A preferred "microorganism having a deregulated isoleucine-valine (*ilv*) pathway" overexpresses acetohydroxyacid synthetase (*e.g.*, acetohydroxyacid synthetase having subunits having the amino acid sequences of SEQ ID NO:32 and SEQ ID NO:34 or acetohydroxyacid synthetase having the amino acid sequence of SEQ ID NO:87), acetohydroxyacid isomeroreductase (having the amino acid sequence of SEQ ID NO:36), or dihydroxyacid dehydratase (having the amino acid

sequence of SEQ ID NO:38) and/or has been transformed with a vector comprising *ilvB*, *ilvN*, *ilvC*, *ilvBN*, *ilvBNC* or *alsS* nucleic acid sequences (SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, nucleotides 1-2246 of SEQ ID NO:58, SEQ ID NO:58 having coding regions from nucleotides 1-1725, 1722-2246 and 2263-3291, or SEQ ID NO:86, respectively) and/or an *ilvD* nucleic acid sequence (SEQ ID NO:37).

#### IV. High Yield Production Methodologies

A particularly preferred embodiment of the present invention is a high yield production method for producing pantothenate comprising culturing a manipulated microorganism under conditions such that pantothenate is produced at a significantly high yield. The phrase "high yield production method", for example, a high yield production method for producing a desired compound (e.g., for producing a panto-compound) includes a method that results in production of the desired compound at a level which is elevated or above what is usual for comparable production methods. Preferably, a high yield production method results in production of the desired compound at a significantly high yield. The phrase "significantly high yield" includes a level of production or yield which is sufficiently elevated or above what is usual for comparable production methods, for example, which is elevated to a level sufficient for commercial production of the desired product (e.g., production of the product at a commercially feasible cost). In one embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 2 g/L. In another embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 10 g/L. In another embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 20 g/L. In yet another embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 30 g/L. In yet another embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 40 g/L. The invention further features a high yield production method for producing a desired compound (e.g., for producing a panto-compound) that involves culturing a manipulated microorganism under conditions such that a sufficiently elevated level of

compound is produced within a commercially desirable period of time. In an exemplary embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 15-20 g/L in 36 hours. In another embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 25-30 g/L in 48 hours. In another embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 35-40 g/L in 72 hours, for example, greater than 37 g/L in 72 hours. In another embodiment, the invention features a high yield production method of producing pantothenate that includes culturing a manipulated microorganism under conditions such that pantothenate is produced at a level greater than 30-40 g/L in 60 hours, for example, greater than 30, 35 or 40 g/L in 60 hours. Values and ranges included and/or intermediate within the ranges set forth herein are also intended to be within the scope of the present invention. For example, pantothenate production at levels of at least 31, 32, 33, 34, 35, 36, 37, 38 and 39 g/L in 60 hours are intended to be included within the range of 30-40 g/L in 60 hours. In another example, ranges of 30-35 g/L or 35-40 g/L are intended to be included within the range of 30-40 g/L in 60 hours. Moreover, the skilled artisan will appreciate that culturing a manipulated microorganism to achieve a production level of, for example, "30-40 g/L in 60 hours" includes culturing the microorganism for additional time periods (e.g., time periods longer than 60 hours), optionally resulting in even higher yields of pantothenate being produced.

25

V. Panto-Compound Production Methodologies Featuring Pantothenate Kinase Mutant Microorganisms

The present invention relates to methods of producing pantothenate using microorganisms engineered to produce high yields of pantothenate as well as other panto-compounds. Cells overproducing pantothenate result in high intracellular pantothenate levels that could overcome the feedback inhibition of pantothenate kinase by CoA, leading to overproduction of CoA. Besides consuming pantothenate, increased synthesis of CoA may cause increased feedback inhibition of the PanB, PanD, PanE or PanC reaction, thereby limiting pantothenate production. Accordingly, a reduction in pantothenate kinase activity may lead to a decrease in CoA levels with resulting increases in PanB, PanD, PanE or PanC activity and pantothenate production.

35

Thus, certain methodologies of the present invention are based, at least in part, on the identification and characterization of the *B. subtilis coaA* gene and the demonstration that the gene is neither essential for *B. subtilis* growth (*i.e.*, deletion of the *coaA* gene from the chromosome of *B. subtilis* is not a lethal event) nor for pantothenate kinase activity in *B. subtilis*. A second pantothenate kinase-encoding gene has been identified and characterized in *B. subtilis*, and is termed "*coaX*". This gene complements an *E. coli* mutant that contains a temperature sensitive pantothenate kinase and is not related by homology to any previously known pantothenate kinase gene.

In one aspect, the methodologies of the invention feature recombinant microorganisms that include the *coaX* gene or that include a mutant *coaX* gene, having reduced pantothenate kinase activity. In one embodiment, the methodologies feature such recombinant microorganisms further having a deregulated pantothenate biosynthetic pathway. In another embodiment, the methodologies feature such recombinant microorganisms further having a deregulated isoleucine-valine (*ilv*) pathway. In a preferred embodiment, the microorganisms belong to the genus *Bacillus* (*e.g.*, *B. subtilis*).

The methodologies of the invention also feature recombinant microorganisms (*e.g.*, microorganisms belong to the genus *Bacillus*, for example, *B. subtilis*) that include the *coaA* gene or that include a mutant *coaA* gene, optionally including a *coaX* gene or mutant thereof, having reduced pantothenate kinase activity. In one embodiment, the methodologies feature such recombinant microorganisms further having a deregulated pantothenate biosynthetic pathway or having a deregulated isoleucine-valine (*ilv*) pathway. Also featured are vectors that include isolated *coaX* or *coaA* genes as well as mutant *coaX* and/or *coaA* genes. Isolated nucleic acid molecules that include isolated *coaX* genes or mutant *coaX* genes are features in addition to isolated CoaX proteins and mutant CoaX proteins.

The above-described nucleic acid molecules (*e.g.*, genes), proteins, vectors, and recombinant microorganisms (*e.g.*, mutant microorganisms), are particularly suited for use in methods of producing panto-compounds and/or methods of enhancing panto-compound production. In one embodiment, the invention features a method for producing a panto-compound (*e.g.*, pantothenate) that includes culturing a pantothenate kinase mutant (*e.g.*, a recombinant microorganism that misexpresses, *e.g.*, is mutated for, pantothenate kinase, as defined herein) under conditions such that panto-compound is produced. In another embodiment, the invention features a method for enhancing production of a panto-compound (*e.g.*, pantothenate) that includes culturing a pantothenate kinase mutant (*e.g.*, a recombinant microorganism that misexpresses, *e.g.*, is mutated for, pantothenate kinase, as defined herein) under conditions such that

production of the panto-compound is produced. As used herein, the term "enhancing" (for example, in the context of the phrase "enhancing production") includes increasing the level or rate of production of panto-compound (e.g., pantothenate) as compared to the level or rate of production in a non-mutant microorganism (e.g., a microorganism having a normal pantothenate kinase gene(s) and/or having normal pantothenate production rates and/or levels.

Preferably, the level of panto-compound produced in methodologies featuring the pantothenate kinase mutants of the present invention is increased by at least 5% as compared to the level produced by a non-mutant (e.g., a recombinant microorganism expressing non-mutated pantothenate kinase). Even more preferably, the level of panto-compound is increased 10% as compared to methodologies featuring non-mutants. Even more preferably, panto-compound levels (e.g., pantothenate levels) are increased 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, are increased 2-fold, 5-fold, 10-fold, 50-fold, 100-fold or more as compared to methodologies featuring non-mutants.

15

#### VI. Additional Mutations Resulting in Enhanced Panto-Compound Production

The methodologies of the present invention further can include, for example in addition to overexpressing or deregulating a pantothenate biosynthetic enzyme and/or an isoleucine-valine biosynthetic enzyme, or in addition to mutating a pantothenate-kinase encoding gene, deleting or mutating an enzyme that catalyzes the conversion of key pantothenate biosynthesis substrates or precursors to unwanted or undesirable products. For example, mutating the *ilvE* gene (Kuramitsu *et al.* (1985) *J. Biochem.* 97:993-999) or a homologue thereof (SEQ ID NO:62 or SEQ ID NO:64), thereby limiting the conversion of  $\alpha$ -ketoisovalerate to valine, in addition to mutating a pantothenate kinase encoding enzyme, is predicted to result in even further enhanced or increased production of panto-compound. Alternatively, mutating the *ansB* gene (Sun and Seflow (1991) *J. Bacteriol.* 173:3831-3845) or a homologue thereof (SEQ ID NO:66), thereby limiting the degradation of aspartate, in addition to mutating a pantothenate kinase encoding enzyme, is predicted to result in even further enhanced or increased production of panto-compound. Alternatively, mutating the *alsD* gene (Renna *et al.* (1993) *J. Bacteriol.* 175:3863-3875) or a homologue thereof (SEQ ID NO:68), thereby limiting the conversion of acetolactate to acetoin, in addition to mutating a pantothenate kinase encoding enzyme, is predicted to result in even further enhanced or increased production of panto-compound. Alternatively, mutating the *avtA* gene encoding alanine-valine transaminase or a homologue thereof, thereby limiting the conversion of  $\alpha$ -ketoisovalerate to valine, in addition to mutating a pantothenate kinase encoding

35

enzyme, is predicted to result in even further enhanced or increased production of panto-compound. Mutating the *avtA* gene can include mutating, for example, an *avtA* gene having the nucleotide sequence of SEQ ID NO:70 (e.g., the *E. coli avtA* gene), or a structural homolog thereof (e.g., a homologue encoding a protein having 30-40%, 40-50%, 50-60%, 60-70%, 70-80%, 80-90%, 90-95% or more identity with the amino acid sequence of SEQ ID NO:71) or a functional homologue (e.g., a gene encoding a structurally unrelated protein having alanine-valine transaminase activity. Such mutations can be accomplished using the methodologies as exemplified in the Examples (e.g., Examples XIII, XV, XVI and XVII).

10 Accordingly, in one embodiment, the invention features a method of producing a panto-compound which includes culturing a microorganism having a mutant pantothenate kinase-encoding gene and which further has a deletion or mutation in an *avtA*, *ilvE*, *ansB*, and/or *alsD* gene, or homologue thereof. In another embodiment, the invention features a method of producing a panto-compound which includes culturing a  
15 microorganism having a mutant pantothenate-kinase encoding gene and a deregulated pantothenate biosynthetic pathway enzyme and which further has a deletion or mutation in an *avtA*, *ilvE*, *ansB*, and/or *alsD* gene, or homologue thereof. In another embodiment, the invention features a method of producing a panto-compound which includes culturing a microorganism having a mutant pantothenate-kinase encoding gene and a  
20 deregulated isoleucine-valine biosynthetic pathway enzyme and which further has a deletion or mutation in an *avtA*, *ilvE*, *ansB*, and/or *alsD* gene, or homologue thereof.

Mutating the *alsD* gene can be particularly useful when accomplished in conjunction with overexpression or deregulation of the *alsS* gene, for example, to prevent carbon (e.g., acetolactate) from being drawn away from the precursor pool  
25 utilized for  $\alpha$ -KIV production. Accordingly, to maximize the contribution of the *als* locus to panto-compound production, it is desirable to disrupt the *alsD* gene in addition to overexpressing the *alsS* gene. To disrupt the *alsD* gene, appropriate fragments of the *als* operon, flanking the *alsD* gene, are amplified by PCR and cloned to provide homology for creating the disruptions. A drug resistance gene, such as the *cat* gene, is  
30 cloned between the flanking DNA fragments in place of the *alsD* gene, and the linearized DNA is transformed into a pantothenate production strain such as PA824, selecting for drug-resistance. To overexpress *alsS*, the *alsS* coding sequence (e.g., an *alsS* coding sequence that has been engineered by PCR for expression) is cloned into an expression vector. Vectors which express *alsS* (or alternatively, vectors which express  
35 *alsS* plus *ilvC*) are introduced into panto-compound production strains (e.g., the pantothenate producing strain PA824).

The methodologies of the present invention further can include, for example in addition to overexpressing or deregulating a pantothenate biosynthetic enzyme and/or an isoleucine-valine biosynthetic enzyme, or in addition to mutating a pantothenate-kinase encoding gene, deleting or mutating an enzyme that catalyzes the conversion of desired panto-compounds to unwanted or undesirable downstream products.

### VII. Isolated Nucleic Acid Molecules and Genes

Another aspect of the present invention features isolated nucleic acid molecules that encode *Bacillus* proteins (e.g., *B. subtilis* proteins), for example, *Bacillus* pantothenate biosynthetic enzymes (e.g., *B. subtilis* pantothenate biosynthetic enzymes) or *Bacillus* valine-isoleucine biosynthetic enzymes (e.g., *B. subtilis* valine-isoleucine biosynthetic enzymes). Also featured are isolated *coaX* and/or *coaA* nucleic acid molecules (e.g., isolated *coaX* and/or *coaA* genes) as well as isolated nucleic acid molecules that include such *coaX* and/or *coaA* nucleic acid molecules or genes.

The term "nucleic acid molecule" includes DNA molecules (e.g., linear, circular, cDNA or chromosomal DNA) and RNA molecules (e.g., tRNA, rRNA, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. The term "isolated" nucleic acid molecule includes a nucleic acid molecule which is free of sequences which naturally flank the nucleic acid molecule (i.e., sequences located at the 5' and 3' ends of the nucleic acid molecule) in the chromosomal DNA of the organism from which the nucleic acid is derived. In various embodiments, an isolated nucleic acid molecule can contain less than about 10 kb, 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb, 0.1 kb, 50 bp, 25 bp or 10 bp of nucleotide sequences which naturally flank the nucleic acid molecule in chromosomal DNA of the microorganism from which the nucleic acid molecule is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular materials when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

The term "gene", as used herein, includes a nucleic acid molecule (e.g., a DNA molecule or segment thereof), for example, a protein or RNA-encoding nucleic acid molecule, that in an organism, is separated from another gene or other genes, by intergenic DNA (i.e., intervening or spacer DNA which naturally flanks the gene and/or separates genes in the chromosomal DNA of the organism). A gene may direct synthesis of an enzyme or other protein molecule (e.g., may comprise coding sequences, for example, a contiguous open reading frame (ORF) which encodes a protein) or may itself be functional in the organism. A gene in an organism, may be clustered in an

operon, as defined herein, said operon being separated from other genes and/or operons by the intergenic DNA. Individual genes contained within an operon may overlap without intergenic DNA between said individual genes. An "isolated gene", as used herein, includes a gene which is essentially free of sequences which naturally flank the gene in the chromosomal DNA of the organism from which the gene is derived (*i.e.*, is free of adjacent coding sequences which encode a second or distinct protein or RNA molecule, adjacent structural sequences or the like) and optionally includes 5' and 3' regulatory sequences, for example promoter sequences and/or terminator sequences. In one embodiment, an isolated gene includes predominantly coding sequences for a protein (*e.g.*, sequences which encode *Bacillus* proteins). In another embodiment, an isolated gene includes coding sequences for a protein (*e.g.*, for a *Bacillus* protein) and adjacent 5' and/or 3' regulatory sequences from the chromosomal DNA of the organism from which the gene is derived (*e.g.*, adjacent 5' and/or 3' *Bacillus* regulatory sequences). Preferably, an isolated gene contains less than about 10 kb, 5 kb, 2 kb, 1 kb, 0.5 kb, 0.2 kb, 0.1 kb, 50 bp, 25 bp or 10 bp of nucleotide sequences which naturally flank the gene in the chromosomal DNA of the organism from which the gene is derived.

In one aspect, the present invention features isolated *panB* nucleic acid sequences or genes, isolated *panC* nucleic acid sequences or genes, isolated *panD* nucleic acid sequences or genes, isolated *panE* nucleic acid sequences or genes, isolated *ilvB*, *ilvN*, *ilvBN* nucleic acid sequences or genes, isolated *alsS* nucleic acid sequences or genes, isolated *ilvC* nucleic acid sequences or genes and/or isolated *ilvD* nucleic acid sequences or genes.

In a preferred embodiment, the nucleic acid or gene is derived from *Bacillus* (*e.g.*, is *Bacillus*-derived). The term "derived from *Bacillus*" or "*Bacillus*-derived" includes a nucleic acid or gene which is naturally found in microorganisms of the genus *Bacillus*. Preferably, the nucleic acid or gene is derived from a microorganism selected from the group consisting of *Bacillus subtilis*, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus firmus*, *Bacillus pantothenicus*, *Bacillus amyloliquefaciens*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus thuringiensis*, and other Group 1 *Bacillus* species, for example, as characterized by 16S rRNA type (Priest, *supra*). In another preferred embodiment, the nucleic acid or gene is derived from *Bacillus brevis* or *Bacillus stearothermophilus*. In another preferred embodiment, the nucleic acid molecules and/or genes of the present invention are derived from a microorganism selected from the group consisting of *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus halodurans*, *Bacillus subtilis*, and *Bacillus pumilus*. In a particularly preferred



embodiment, the nucleic acid or gene is derived from *Bacillus subtilis* (e.g., is *Bacillus subtilis*-derived). The term "derived from *Bacillus subtilis*" or "*Bacillus subtilis*-derived" includes a nucleic acid or gene which is naturally found in *Bacillus subtilis*. In yet another preferred embodiment, the nucleic acid or gene is a *Bacillus* gene  
5 homologue (e.g., is derived from a species distinct from *Bacillus* but having significant homology to a *Bacillus* gene of the present invention, for example, a *Bacillus pan* gene or *Bacillus ilv* gene).

Included within the scope of the present invention are bacterial-derived nucleic acid molecules or genes and/or *Bacillus*-derived nucleic acid molecules or genes (e.g.,  
10 *B. subtilis*-derived nucleic acid molecules or genes), for example, the genes identified by the present inventors, for example, *Bacillus* or *B. subtilis coaX* genes, *coaA* genes, *pan* genes and/or *ilv* genes. Further included within the scope of the present invention are bacterial-derived nucleic acid molecules or genes and/or *Bacillus*-derived nucleic acid molecules or genes (e.g., *B. subtilis*-derived nucleic acid molecules or genes) (e.g., *B.*  
15 *subtilis* nucleic acid molecules or genes) which differ from naturally-occurring bacterial and/or *Bacillus* nucleic acid molecules or genes (e.g., *B. subtilis* nucleic acid molecules or genes), for example, nucleic acid molecules or genes which have nucleic acids that are substituted, inserted or deleted, but which encode proteins substantially similar to the naturally-occurring gene products of the present invention. In one embodiment, an  
20 isolated nucleic acid molecule comprises at least one of the nucleotide sequences set forth as SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO 31, SEQ ID NO:33, SEQ ID NO:86, SEQ ID NO:35 or SEQ ID NO:37. In another preferred embodiment, an isolated nucleic acid molecule comprises at least two, three or four of the nucleotide sequences set forth as SEQ ID NO:23, SEQ ID NO:25, SEQ ID  
25 NO:27, SEQ ID NO:29, SEQ ID NO 31, SEQ ID NO:33, SEQ ID NO:88, SEQ ID NO:35 or SEQ ID NO:37. For example, a preferred isolated nucleic acid molecule of the present invention can include the nucleotide sequences of SEQ ID NO:23, SEQ ID NO:25 and SEQ ID NO:27, preferably linked such that the proteins encoded by the nucleotide sequences of SEQ ID NO:23, SEQ ID NO:25 and SEQ ID NO:27 are each  
30 produced when the isolated nucleic acid molecule is expressed in a microorganism (e.g., SEQ ID NO:59). In another example, a preferred isolated nucleic acid molecule of the present invention can include the nucleotide sequences of SEQ ID NO:31 and SEQ ID NO:33, preferably linked such that the proteins encoded by the nucleotide sequences of SEQ ID NO:31 and SEQ ID NO:33 are each produced when the isolated nucleic acid  
35 molecule is expressed in a microorganism (e.g., nucleotides 1-2246 of SEQ ID NO:58). In another example, a preferred isolated nucleic acid molecule of the present invention can include the nucleotide sequence of SEQ ID NO:86. In another example, a preferred

isolated nucleic acid molecule of the present invention can include the nucleotide sequences of SEQ ID NO:31, SEQ ID NO:33 and SEQ ID NO:35, preferably linked such that the proteins encoded by the nucleotide sequences of SEQ ID NO:31, SEQ ID NO:33 and SEQ ID NO:35 are each produced when the isolated nucleic acid molecule is expressed in a microorganism (*e.g.*, SEQ ID NO:58).

In another embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 60-65%, preferably at least about 70-75%, more preferable at least about 80-85%, and even more preferably at least about 90-95% or more identical to a nucleotide sequence set forth as SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO 31, SEQ ID NO:33, SEQ ID NO:88, SEQ ID NO:35 or SEQ ID NO:37. In another embodiment, an isolated nucleic acid molecule hybridizes under stringent conditions to a nucleic acid molecule having a nucleotide sequence set forth as SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO 31, SEQ ID NO:33, SEQ ID NO:88, SEQ ID NO:35 or SEQ ID NO:37. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent (*e.g.* high stringency) hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO 31, SEQ ID NO:33, SEQ ID NO:88, SEQ ID NO:35 or SEQ ID NO:37 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature.

A nucleic acid molecule of the present invention (*e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO 31, SEQ ID NO:33, SEQ ID NO:88, SEQ ID NO:35 or SEQ ID NO:37 can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) or can be isolated by the polymerase chain reaction using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO 31, SEQ ID NO:33, SEQ ID NO:88, SEQ ID NO:35 or SEQ ID NO:37. A nucleic acid of the invention can be

amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide  
5 sequence shown in SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:33, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:88, SEQ ID NO:35.

Additional *panC* nucleic acid sequences include those that comprise the nucleotide sequence of SEQ ID NO:25, encode a homologue of the polypeptide having the amino acid sequence set forth in SEQ ID NO:26 (e.g., encode a polypeptide having  
10 at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or more identity to the polypeptide having the amino acid sequence as set forth in SEQ ID NO:26 and a substantially identical activity as said polypeptide), hybridize under stringent conditions to all or a portion of a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:25 or to all or a portion of a nucleic acid molecule that encodes a polypeptide having the  
15 amino acid sequence of SEQ ID NO:26, or are complementary to a *panC* nucleotide sequence as set forth herein.

Additional *panD* nucleic acid sequences include those that comprise the nucleotide sequence of SEQ ID NO:27, encode a homologue of the polypeptide having the amino acid sequence set forth in SEQ ID NO:28 (e.g., encode a polypeptide having  
20 at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or more identity to the polypeptide having the amino acid sequence as set forth in SEQ ID NO:28 and a substantially identical activity as said polypeptide), hybridize under stringent conditions to all or a portion of a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:27 or to all or a portion of a nucleic acid molecule that encodes a polypeptide having the  
25 amino acid sequence of SEQ ID NO:28, or are complementary to a *panD* nucleotide sequence as set forth herein.

Additional *panE* nucleic acid sequences include those that comprise the nucleotide sequence of SEQ ID NO:29, encode a homologue of the polypeptide having the amino acid sequence set forth in SEQ ID NO:30 (e.g., encode a polypeptide having  
30 at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or more identity to the polypeptide having the amino acid sequence as set forth in SEQ ID NO:30 and a substantially identical activity as said polypeptide), hybridize under stringent conditions to all or a portion of a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:29 or to all or a portion of a nucleic acid molecule that encodes a polypeptide having the  
35 amino acid sequence of SEQ ID NO:30, or are complementary to a *panE* nucleotide sequence as set forth herein.

Additional *ilvB* nucleic acid sequences are those that comprise the nucleotide sequence of SEQ ID NO:31, encode a homologue of the polypeptide having the amino acid sequence set forth in SEQ ID NO:32 (*e.g.*, encode a polypeptide having at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or more identity to the polypeptide having the amino acid sequence as set forth in SEQ ID NO:32 and a substantially identical activity as said polypeptide), hybridize under stringent conditions to all or a portion of a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:31 or to all or a portion of a nucleic acid molecule that encodes a polypeptide having the amino acid sequence of SEQ ID NO:32, or are complementary to an *ilvB* nucleotide sequence as set forth herein.

Additional *ilvN* nucleic acid sequences are those that comprise the nucleotide sequence of SEQ ID NO:33, encode a homologue of the polypeptide having the amino acid sequence set forth in SEQ ID NO:34 (*e.g.*, encode a polypeptide having at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or more identity to the polypeptide having the amino acid sequence as set forth in SEQ ID NO:34 and a substantially identical activity as said polypeptide), hybridize under stringent conditions to all or a portion of a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:33 or to all or a portion of a nucleic acid molecule that encodes a polypeptide having the amino acid sequence of SEQ ID NO:34, or are complementary to an *ilvN* nucleotide sequence as set forth herein.

Additional *ilvC* nucleic acid sequences include those that comprise the nucleotide sequence of SEQ ID NO:35, encode a homologue of the polypeptide having the amino acid sequence set forth in SEQ ID NO:36 (*e.g.*, encode a polypeptide having at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or more identity to the polypeptide having the amino acid sequence as set forth in SEQ ID NO:36 and a substantially identical activity as said polypeptide), hybridize under stringent conditions to all or a portion of a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:35 or to all or a portion of a nucleic acid molecule that encodes a polypeptide having the amino acid sequence of SEQ ID NO:36, or are complementary to an *ilvC* nucleotide sequence as set forth herein.

Additional *ilvD* nucleic acid sequences include those that comprise the nucleotide sequence of SEQ ID NO:37, encode a homologue of the polypeptide having the amino acid sequence set forth in SEQ ID NO:38 (*e.g.*, encode a polypeptide having at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or more identity to the polypeptide having the amino acid sequence as set forth in SEQ ID NO:38 and a substantially identical activity as said polypeptide), hybridize under stringent conditions to all or a portion of a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:37 or

to all or a portion of a nucleic acid molecule that encodes a polypeptide having the amino acid sequence of SEQ ID NO:38, or are complementary to an *ilvD* nucleotide sequence as set forth herein.

Additional *alsS* nucleic acid sequences include those that comprise the  
5 nucleotide sequence of SEQ ID NO:86, encode a homologue of the polypeptide having the amino acid sequence set forth in SEQ ID NO:87 (*e.g.*, encode a polypeptide having at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or more identity to the polypeptide having the amino acid sequence as set forth in SEQ ID NO:87 and a substantially identical activity as said polypeptide), hybridize under stringent conditions to all or a  
10 portion of a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:86 or to all or a portion of a nucleic acid molecule that encodes a polypeptide having the amino acid sequence of SEQ ID NO:87, or are complementary to an *alsS* nucleotide sequence as set forth herein.

In another embodiment, an isolated nucleic acid molecule is or includes a *coaX*  
15 gene, or portion or fragment thereof. In one embodiment, an isolated *coaX* nucleic acid molecule or gene comprises the nucleotide sequence as set forth in SEQ ID NO:19 (*e.g.*, comprises the *B. subtilis coaX* nucleotide sequence). In another embodiment, an isolated *coaX* nucleic acid molecule or gene comprises a nucleotide sequence that encodes the amino acid sequence as set forth in SEQ ID NO:9 (*e.g.*, encodes the *B.*  
20 *subtilis* CoaX amino acid sequence). In yet another embodiment, an isolated *coaX* nucleic acid molecule or gene encodes a homologue of the CoaX protein having the amino acid sequence of SEQ ID NO:9. As used herein, the term "homologue" includes a protein or polypeptide sharing at least about 30-35%, preferably at least about 35-40%, more preferably at least about 40-50%, and even more preferably at least about 60%,  
25 70%, 80%, 90% or more identity with the amino acid sequence of a wild-type protein or polypeptide described herein and having a substantially equivalent functional or biological activity as said wild-type protein or polypeptide. For example, a CoaX homologue shares at least about 30-35%, preferably at least about 35-40%, more preferably at least about 40-50%, and even more preferably at least about 60%, 70%,  
30 80%, 90% or more identity with the protein having the amino acid sequence set forth as SEQ ID NO:9 and has a substantially equivalent functional or biological activity (*i.e.*, is a functional equivalent) of the protein having the amino acid sequence set forth as SEQ ID NO:9 (*e.g.*, has a substantially equivalent pantothenate kinase activity). In a preferred embodiment, an isolated *coaX* nucleic acid molecule or gene comprises a  
35 nucleotide sequence that encodes a polypeptide as set forth in any one of SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID

NO:74 or SEQ ID NO:75. In another embodiment, an isolated *coaX* nucleic acid molecule hybridizes to all or a portion of a nucleic acid molecule having the nucleotide sequence set forth in SEQ ID NO:19 or hybridizes to all or a portion of a nucleic acid molecule having a nucleotide sequence that encodes a polypeptide having the amino acid sequence of any of SEQ ID NOs:7-18, 74 or 75. Such hybridization conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1X SSPE is 0.15 M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25 mM EDTA, pH 7.4) can be substituted for SSC (1X SSC is 0.15 M NaCl and 15 mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature ( $T_m$ ) of the hybrid, where  $T_m$  is determined according to the following equations. For hybrids less than 18 base pairs in length,  $T_m(^{\circ}\text{C}) = 2(\# \text{ of A} + \text{T bases}) + 4(\# \text{ of G} + \text{C bases})$ . For hybrids between 18 and 49 base pairs in length,  $T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G} + \text{C}) - (600/\text{N})$ , where N is the number of bases in the hybrid, and  $[\text{Na}^+]$  is the concentration of sodium ions in the hybridization buffer ( $[\text{Na}^+]$  for 1X SSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of

stringent hybridization conditions is hybridization in 0.25-0.5M NaH<sub>2</sub>PO<sub>4</sub>, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH<sub>2</sub>PO<sub>4</sub>, 1% SDS at 65°C, see e.g., Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or, alternatively, 0.2X SSC, 1% SDS). In another preferred embodiment, an isolated

5 nucleic acid molecule comprises a nucleotide sequence that is complementary to a *coaX* nucleotide sequence as set forth herein (e.g., is the full complement of the nucleotide sequence set forth as SEQ ID NO:19).

In another preferred embodiment, an isolated nucleic acid molecule is or includes a *coaA* gene, for example, a *Bacillus* (e.g., *B. subtilis*) *coaA* gene, or portion or

10 fragment thereof. Exemplary isolated *coaA* nucleic acid molecules and/or genes include (1) an isolated *coaA* nucleic acid molecule or gene comprising the nucleotide sequence as set forth in any one of SEQ ID NOs:20-22; (2) an isolated *coaA* nucleic acid molecule or gene comprising a nucleotide sequence that encodes the amino acid sequence as set forth in SEQ ID NO:3; (3) an isolated *coaA* nucleic acid molecule or gene comprising a

15 nucleotide sequence which encodes a CoaA homologue (e.g., a polypeptide having an amino acid sequence at least about 30-35%, preferably at least about 35-40%, more preferably at least about 40-50%, and even more preferably at least about 60%, 70%, 80%, 90% or more identical to the amino acid sequence set forth as SEQ ID NO:3 and having a substantially equivalent enzymatic activity; (4) an isolated *coaA* nucleic acid

20 molecule or gene comprising a nucleotide sequence that encodes a polypeptide as set forth in any one of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5 or SEQ ID NO:6; (5) an isolated nucleic acid molecule that hybridizes under stringent conditions to all or a portion of a nucleic acid molecule having the nucleotide sequence set forth in SEQ ID NO:20, SEQ ID NO:21 or SEQ ID NO:22 or hybridizes to all or a

25 portion of a nucleic acid molecule having a nucleotide sequence that encodes a polypeptide having the amino acid sequence of SEQ ID NO:3; and (6) an isolated nucleic acid molecule comprising a nucleotide sequence that is complementary to a *coaA* nucleotide sequence as set forth herein (e.g., is the full complement of the nucleotide sequence set forth in SEQ ID NO:20, SEQ ID NO:21 or SEQ ID NO:22).

30 A nucleic acid molecule of the present invention (e.g., a *coaX* nucleic acid molecule or gene or a *coaA* nucleic acid molecule or gene), can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, nucleic acid molecules can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) or can be isolated

35 by the polymerase chain reaction using synthetic oligonucleotide primers designed

based upon the *coaX* or *coaA* nucleotide sequences set forth herein, or flanking sequences thereof. A nucleic acid of the invention (e.g., a *coaX* nucleic acid molecule or gene or a *coaA* nucleic acid molecule or gene), can be amplified using cDNA, mRNA or alternatively, chromosomal DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques.

Yet another embodiment of the present invention features mutant *coaX* and *coaA* nucleic acid molecules or genes. The phrase "mutant nucleic acid molecule" or "mutant gene" as used herein, includes a nucleic acid molecule or gene having a nucleotide sequence which includes at least one alteration (e.g., substitution, insertion, deletion) such that the polypeptide or protein that may be encoded by said mutant exhibits an activity that differs from the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene. Preferably, a mutant nucleic acid molecule or mutant gene (e.g., a mutant *coaA* or *coaX* gene) encodes a polypeptide or protein having a reduced activity (e.g., having a reduced pantothenate kinase activity) as compared to the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene, for example, when assayed under similar conditions (e.g., assayed in microorganisms cultured at the same temperature). A mutant gene also can encode no polypeptide or have a reduced level of production of the wild-type polypeptide.

As used herein, a "reduced activity" or "reduced enzymatic activity" is one that is at least 5% less than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene, preferably at least 5-10% less, more preferably at least 10-25% less and even more preferably at least 25-50%, 50-75% or 75-100% less than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene. Ranges intermediate to the above-recited values, e.g., 75-85%, 85-90%, 90-95%, are also intended to be encompassed by the present invention. As used herein, a "reduced activity" or "reduced enzymatic activity" also includes an activity that has been deleted or "knocked out" (e.g., approximately 100% less activity than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene). Activity can be determined according to any well accepted assay for measuring activity of a particular protein of interest. Activity can be measured or assayed directly, for example, measuring an activity of a protein isolated or purified from a cell. Alternatively, an activity can be measured or assayed within a cell or in an extracellular medium. For example, assaying for a mutant *coaA* gene or a mutant *coaX* gene (i.e., said mutant encoding a reduced pantothenate kinase activity) can be accomplished by expressing the mutated gene in a microorganism, for example, a mutant microorganism which expresses pantothenate kinase in a temperature-sensitive manner, assaying the mutant gene for the ability to complement a temperature sensitive (Ts) mutant for



pantothenate kinase activity. A *coaX* mutant gene or *coaA* mutant gene that encodes a "reduced pantothenate kinase activity" is one that complements the Ts mutant less effectively than, for example, a corresponding wild-type *coaX* gene or *coaA* gene.

It will be appreciated by the skilled artisan that even a single substitution in a nucleic acid or gene sequence (e.g., a base substitution that encodes an amino acid change in the corresponding amino acid sequence) can dramatically affect the activity of an encoded polypeptide or protein as compared to the corresponding wild-type polypeptide or protein. A mutant nucleic acid or mutant gene (e.g., encoding a mutant polypeptide or protein), as defined herein, is readily distinguishable from a nucleic acid or gene encoding a protein homologue, as described above, in that a mutant nucleic acid or mutant gene encodes a protein or polypeptide having an altered activity, optionally observable as a different or distinct phenotype in a microorganism expressing said mutant gene or nucleic acid or producing said mutant protein or polypeptide (i.e., a mutant microorganism) as compared to a corresponding microorganism expressing the wild-type gene or nucleic acid or producing said mutant protein or polypeptide. By contrast, a protein homologue has an identical or substantially similar activity, optionally phenotypically indiscernable when produced in a microorganism, as compared to a corresponding microorganism expressing the wild-type gene or nucleic acid. Accordingly it is not, for example, the degree of sequence identity between nucleic acid molecules, genes, protein or polypeptides that serves to distinguish between homologues and mutants, rather it is the activity of the encoded protein or polypeptide that distinguishes between homologues and mutants: homologues having, for example, low (e.g., 30-50% sequence identity) sequence identity yet having substantially equivalent functional activities, and mutants, for example sharing 99% sequence identity yet having dramatically different or altered functional activities. Exemplary homologues are set forth in Figure 20 (i.e., CoaA homologues) and in Figure 23 (i.e., CoaX homologues). Exemplary mutants are described in Examples XV and XVIII herein.

#### VIII. Recombinant Nucleic Acid Molecules and Vectors

The present invention further features recombinant nucleic acid molecules (e.g., recombinant DNA molecules) that include nucleic acid molecules and/or genes described herein (e.g., isolated nucleic acid molecules and/or genes), preferably *Bacillus* genes, more preferably *Bacillus subtilis* genes, even more preferably *Bacillus subtilis* pantothenate kinase genes (e.g., *coaX* genes or *coaA* genes), pantothenate biosynthetic genes (e.g., genes encoding pantothenate biosynthetic enzymes, for example, *panB* genes encoding ketopantoate hydroxymethyltransferase, *panE* genes encoding

ketopantoate reductase, *panC* genes encoding pantothenate synthetase, and/or *panD* genes encoding aspartate- $\alpha$ -decarboxylase) and/or isoleucine-valine (*ilv*) biosynthetic genes (e.g., *ilvBN* or *alsS* genes encoding acetohydroxyacid synthetase, *ilvC* genes encoding acetohydroxyacid isomeroreductase and/or *ilvD* genes encoding dihydroxyacid dehydratase).

The present invention further features vectors (e.g., recombinant vectors) that include nucleic acid molecules (e.g., isolated or recombinant nucleic acid molecules and/or genes) described herein. In particular, recombinant vectors are featured that include nucleic acid sequences that encode bacterial gene products as described herein, preferably *Bacillus* gene products, more preferably *Bacillus subtilis* gene products, even more preferably *Bacillus subtilis* pantothenate biosynthetic gene products (e.g., pantothenate biosynthetic enzymes, for example, ketopantoate hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase, and/or aspartate- $\alpha$ -decarboxylase) and/or isoleucine-valine biosynthetic gene products (e.g., acetohydroxyacid synthetase, acetohydroxyacid isomeroreductase and/or dihydroxyacid dehydratase).

The term "recombinant nucleic acid molecule" includes a nucleic acid molecule (e.g., a DNA molecule) that has been altered, modified or engineered such that it differs in nucleotide sequence from the native or natural nucleic acid molecule from which the recombinant nucleic acid molecule was derived (e.g., by addition, deletion or substitution of one or more nucleotides). Preferably, a recombinant nucleic acid molecule (e.g., a recombinant DNA molecule) includes an isolated nucleic acid molecule or gene of the present invention (e.g., an isolated *coaX*, *coaA*, *pan* or *ilv* gene) operably linked to regulatory sequences.

The term "recombinant vector" includes a vector (e.g., plasmid, phage, phasmid, virus, cosmid or other purified nucleic acid vector) that has been altered, modified or engineered such that it contains greater, fewer or different nucleic acid sequences than those included in the native or natural nucleic acid molecule from which the recombinant vector was derived. Preferably, the recombinant vector includes a *coaX*, *coaA*, *pan* or *ilv* gene or recombinant nucleic acid molecule including such *coaX*, *coaA*, *pan* or *ilv* gene, operably linked to regulatory sequences, for example, promoter sequences, terminator sequences and/or artificial ribosome binding sites (RBSs), as defined herein.

The phrase "operably linked to regulatory sequence(s)" means that the nucleotide sequence of the nucleic acid molecule or gene of interest is linked to the regulatory sequence(s) in a manner which allows for expression (e.g., enhanced, increased, constitutive, basal, attenuated, decreased or repressed expression) of the

nucleotide sequence, preferably expression of a gene product encoded by the nucleotide sequence (e.g., when the recombinant nucleic acid molecule is included in a recombinant vector, as defined herein, and is introduced into a microorganism).

The term "regulatory sequence" includes nucleic acid sequences which affect (e.g., modulate or regulate) expression of other nucleic acid sequences. In one embodiment, a regulatory sequence is included in a recombinant nucleic acid molecule or recombinant vector in a similar or identical position and/or orientation relative to a particular gene of interest as is observed for the regulatory sequence and gene of interest as it appears in nature, e.g., in a native position and/or orientation. For example, a gene of interest can be included in a recombinant nucleic acid molecule or recombinant vector operably linked to a regulatory sequence which accompanies or is adjacent to the gene of interest in the natural organism (e.g., operably linked to "native" regulatory sequences, for example, to the "native" promoter). Alternatively, a gene of interest can be included in a recombinant nucleic acid molecule or recombinant vector operably linked to a regulatory sequence which accompanies or is adjacent to another (e.g., a different) gene in the natural organism. Alternatively, a gene of interest can be included in a recombinant nucleic acid molecule or recombinant vector operably linked to a regulatory sequence from another organism. For example, regulatory sequences from other microbes (e.g., other bacterial regulatory sequences, bacteriophage regulatory sequences and the like) can be operably linked to a particular gene of interest.

In one embodiment, a regulatory sequence is a non-native or non-naturally-occurring sequence (e.g., a sequence which has been modified, mutated, substituted, derivatized, deleted including sequences which are chemically synthesized). Preferred regulatory sequences include promoters, enhancers, termination signals, anti-termination signals and other expression control elements (e.g., sequences to which repressors or inducers bind and/or binding sites for transcriptional and/or translational regulatory proteins, for example, in the transcribed mRNA). Such regulatory sequences are described, for example, in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in a microorganism (e.g., constitutive promoters and strong constitutive promoters), those which direct inducible expression of a nucleotide sequence in a microorganism (e.g., inducible promoters, for example, xylose inducible promoters) and those which attenuate or repress expression of a nucleotide sequence in a microorganism (e.g., attenuation signals or repressor sequences). It is also within the scope of the present invention to regulate expression of a gene of interest by removing or deleting regulatory sequences. For

example, sequences involved in the negative regulation of transcription can be removed such that expression of a gene of interest is enhanced.

In one embodiment, a recombinant nucleic acid molecule or recombinant vector of the present invention includes a nucleic acid sequence or gene that encodes at least one bacterial gene product (e.g., a pantothenate biosynthetic enzyme, an isoleucine-valine biosynthetic enzyme, or a CoaA biosynthetic enzyme, for example CoaA or CoaX) operably linked to a promoter or promoter sequence. Preferred promoters of the present invention include *Bacillus* promoters and/or bacteriophage promoters (e.g., bacteriophage which infect *Bacillus*). In one embodiment, a promoter is a *Bacillus* promoter, preferably a strong *Bacillus* promoter (e.g., a promoter associated with a biochemical housekeeping gene in *Bacillus* or a promoter associated with a glycolytic pathway gene in *Bacillus*). In another embodiment, a promoter is a bacteriophage promoter. In a preferred embodiment, the promoter is from the bacteriophage SP01. In a particularly preferred embodiment, a promoter is selected from the group consisting of  $P_{15}$ ,  $P_{26}$  or  $P_{veg}$ , for example, the promoters set forth in SEQ ID NO:39, SEQ ID NO:40 or SEQ ID NO:41. Additional preferred promoters include *tef* (the translational elongation factor (TEF) promoter) and *pyc* (the pyruvate carboxylase (PYC) promoter), which promote high level expression in *Bacillus* (e.g., *Bacillus subtilis*). Additional preferred promoters, for example, for use in Gram positive microorganisms include, but are not limited to, the *amyE* promoter or phage SP02 promoters. Additional preferred promoters, for example, for use in Gram negative microorganisms include, but are not limited to *tac*, *trp*, *tet*, *trp-tet*, *lpp*, *lac*, *lpp-lac*, *lacIq*, *T7*, *T5*, *T3*, *gal*, *trc*, *ara*, *SP6*,  $\lambda$ - $P_R$  or  $\lambda$ - $P_L$ .

In another embodiment, a recombinant nucleic acid molecule or recombinant vector of the present invention includes a terminator sequence or terminator sequences (e.g., transcription terminator sequences). The term "terminator sequences" includes regulatory sequences which serve to terminate transcription of a gene. Terminator sequences (or tandem transcription terminators) can further serve to stabilize mRNA (e.g., by adding structure to mRNA), for example, against nucleases.

In yet another embodiment, a recombinant nucleic acid molecule or recombinant vector of the present invention includes sequences which allow for detection of the vector containing said sequences (i.e., detectable and/or selectable markers), for example, sequences that overcome auxotrophic mutations, for example, *ura3* or *ilvE*, fluorescent markers, and/or colorimetric markers (e.g., *lacZ*/ $\beta$ -galactosidase), and/or antibiotic resistance genes (e.g., *amp* or *tet*).

In yet another embodiment, a recombinant nucleic acid molecule or recombinant vector of the present invention includes an artificial ribosome binding site (RBS). The term "artificial ribosome binding site (RBS)" includes a site within an mRNA molecule (e.g., coded within DNA) to which a ribosome binds (e.g., to initiate translation) which differs from a native RBS (e.g., a RBS found in a naturally-occurring gene) by at least one nucleotide. Preferred artificial RBSs include about 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, 17-18, 19-20, 21-22, 23-24, 25-26, 27-28, 29-30 or more nucleotides of which about 1-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-15 or more differ from the native RBS (e.g., the native RBS of a gene of interest). Preferably, nucleotides which differ are substituted such that they are identical to one or more nucleotides of an ideal RBS (e.g., SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47 or SEQ ID NO:48), when optimally aligned for comparisons. Artificial RBSs can be used to replace the naturally-occurring or native RBS associated with a particular gene. Artificial RBSs preferably increase translation of a particular gene. Preferred artificial RBSs (e.g., RBSs for increasing the translation of *panB*, for example, of *B. subtilis panB*) are depicted in Table 1A (e.g., SEQ ID NO:49 and SEQ ID NO:50).

Table 1A: Preferred *panB* Ribosome Binding Sites

	10	20	
	-----AGAAAGGAGGTGA		ideal RBS (SEQ ID NO:44)
	CCCTCT-AG-AAGGAGGAGAAAACATG		RBS1 (SEQ ID NO:49)
	CCCTCT-AG--AGGAGGAGAAAACATG		RBS2 (SEQ ID NO:50)
25	TAAACAT-G--AGGAGGAGAAAACATG		<i>panB</i> native RBS (SEQ ID NO:42)

Additional preferred artificial RBSs (e.g., RBSs for increasing the translation of *panD*, for example, of *B. subtilis panD*) are depicted in Table 1B (e.g., SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53 and SEQ ID NO:54).

Table 1B: Preferred *panD* Ribosome Binding Sites

	10	20	
35	CTAGAAAAGGAGGAATTTAAATG		<i>panD</i> 423 RBS (SEQ ID NO:88)
	TTAAGAAAGGAGGTGANNNNATG		ideal RBS (SEQ ID NO:45)
	TTAGAAAGGAGGATTTAAATATG		new design A (SEQ ID NO:51)
	TTAGAAAGGAGGTTTAATTAATG		new design B (SEQ ID NO:52)
40	TTAGAAAGGAGGTGATTTAAATG		new design C1 (SEQ ID NO:53)
	TTAGAAAGGAGGTGTTTAAAATG		new design C2 (SEQ ID NO:54)
	TTAGAAAGGAGGTGANNNNATG		ideal RBS (SEQ ID NO:46)

Additional preferred artificial RBSs (e.g., RBSs for increasing the translation of *panD*, for example, of *B. subtilis panD*) are depicted in Table 1C (e.g., SEQ ID NO:55, SEQ ID NO:56 and SEQ ID NO:57). The predicted amino acid sequence at the C-terminus of the PanC protein is shown. The start codon for PanD translation is underlined.

Table 1C: Additional Preferred *panD* Ribosome Binding Sites

	10	20	
10	---	---A GAA AGG AGG TGA NNN NNN N <u>ATG</u>	ideal RBS (SEQ ID NO:47)
	ATT CGA GAA ATG GAG AGA ATA TAA T <u>ATG</u>		native <i>panD</i> RBS (SEQ ID NO:43)
	Ile Arg Glu Met Glu Arg Ile *	Met	SEQ ID NO:89
15	---	---A GAA AGG AGG TGA NNN NNN N <u>ATG</u>	ideal RBS (SEQ ID NO:47)
	ATT CGA GAA AGG AGG TGA ATA TAA T <u>ATG</u>		NDI (SEQ ID NO:55)
	Ile Arg Glu Arg Arg *	Met	SEQ ID NO:90
20	ATT CGA GAA AGG AGG TGA ATA ATA - <u>ATG</u>		NDII (SEQ ID NO:56)
	Ile Arg Glu Arg Arg *	Met	SEQ ID NO:90
	ATT CGT AGA AAG GAG GTG AAT TAA T <u>ATG</u>		NDIII (SEQ ID NO:57)
25	Ile Arg Arg Lys Glu Val Asn *	Met	SEQ ID NO:91
	---	---AGA AAG GAG GTG ANN NNN N <u>ATG</u>	ideal RBS (SEQ ID NO:48)

Accordingly, in one embodiment, a vector of the present invention includes an artificial RBS as set forth in SEQ ID NO:49 or SEQ ID NO:50. In another embodiment, a vector of the present invention includes an artificial RBS as set forth in SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53 or SEQ ID NO:54. In yet another embodiment, a vector of the present invention includes an artificial RBS as set forth in SEQ ID NO:55, SEQ ID NO:56 or SEQ ID NO:57.

In another embodiment, a recombinant vector of the present invention includes sequences that enhance replication in bacteria (e.g., replication-enhancing sequences). In one embodiment, replication-enhancing sequences are derived from *E. coli*. In another embodiment, replication-enhancing sequences are derived from pBR322 (e.g., sequences included within the pBR322 derived portion of any of the pAN vectors as set forth in the Figures, i.e., the Not I-Not I sequences from about 5.0 kB to 9.0 kB of the vector depicted in Figure 3A).

In yet another embodiment, a recombinant vector of the present invention includes antibiotic resistance genes. The term "antibiotic resistance genes" includes sequences which promote or confer resistance to antibiotics on the host organism (e.g., *Bacillus*). In one embodiment, the antibiotic resistance genes are selected from the group consisting of *cat* (chloramphenicol resistance) genes, *tet* (tetracycline resistance) genes, *erm* (erythromycin resistance) genes, *neo* (neomycin resistance) genes and *spec* (spectinomycin resistance) genes. Recombinant vectors of the present invention can further include homologous recombination sequences (e.g., sequences designed to allow recombination of the gene of interest into the chromosome of the host organism). For example, *amyE* sequences can be used as homology targets for recombination into the host chromosome.

Preferred vectors of the present invention include, but are not limited to, vectors set forth in Figures 2-15, 17, 19, 22, 25 and 26. It will further be appreciated by one of skill in the art that the design of a vector can be tailored depending on such factors as the choice of microorganism to be genetically engineered, the level of expression of gene product desired and the like.

#### IX. Isolated Proteins

Another aspect of the present invention features isolated proteins (e.g., isolated pantothenate biosynthetic enzymes and/or valine-isoleucine biosynthetic enzymes and/or isolated CoA biosynthetic enzymes, for example isolated CoaA or CoaX). In one embodiment, proteins (e.g., isolated pantothenate biosynthetic enzymes and/or valine-isoleucine biosynthetic enzymes and/or isolated CoaA biosynthetic enzymes, for example isolated CoaA or CoaX) are produced by recombinant DNA techniques and can be isolated from microorganisms of the present invention by an appropriate purification scheme using standard protein purification techniques. In another embodiment, proteins are synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein (e.g., an isolated or purified biosynthetic enzyme) is substantially free of cellular material or other contaminating proteins from the microorganism from which the protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. In one embodiment, an isolated or purified protein has less than about 30% (by dry weight) of contaminating protein or chemicals, more preferably less than about 20% of contaminating protein or chemicals, still more preferably less than about 10% of contaminating protein or chemicals, and most preferably less than about 5% contaminating protein or chemicals.

In a preferred embodiment, the protein or gene product is derived from *Bacillus* (e.g., is *Bacillus*-derived). The term "derived from *Bacillus*" or "*Bacillus*-derived" includes a protein or gene product which is encoded by a *Bacillus* gene. Preferably, the gene product is derived from a microorganism selected from the group consisting of

5 *Bacillus subtilis*, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus firmus*, *Bacillus pantothenicus*, *Bacillus amyloliquefaciens*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus thuringiensis*, and other Group 1 *Bacillus* species, for example, as characterized by 16S rRNA type (Priest, *supra*). In another preferred embodiment, the protein or gene

10 product is derived from *Bacillus brevis* or *Bacillus stearothermophilus*. In another preferred embodiment, the protein or gene product is derived from a microorganism selected from the group consisting of *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus halodurans*, *Bacillus subtilis*, and *Bacillus pumilus*. In a particularly preferred embodiment, the protein or gene product is derived from *Bacillus subtilis* (e.g., is

15 *Bacillus subtilis*-derived). The term "derived from *Bacillus subtilis*" or "*Bacillus subtilis*-derived" includes a protein or gene product which is encoded by a *Bacillus subtilis* gene. In yet another preferred embodiment, the protein or gene product is encoded by a *Bacillus* gene homologue (e.g., a gene derived from a species distinct from *Bacillus* but having significant homology to a *Bacillus* gene of the present invention, for

20 example, a *Bacillus pan* gene or *Bacillus ilv* gene).

Included within the scope of the present invention are bacterial-derived proteins or gene products and/or *Bacillus*-derived proteins or gene products (e.g., *B. subtilis*-derived gene products) that are encoded by naturally-occurring bacterial and/or *Bacillus* genes (e.g., *B. subtilis* genes), for example, the genes identified by the present inventors,

25 for example, *Bacillus* or *B. subtilis coaX* genes, *coaA* genes, *pan* genes and/or *ilv* genes. Further included within the scope of the present invention are bacterial-derived proteins or gene products and/or *Bacillus*-derived proteins or gene products (e.g., *B. subtilis*-derived gene products) that are encoded bacterial and/or *Bacillus* genes (e.g., *B. subtilis* genes) which differ from naturally-occurring bacterial and/or *Bacillus* genes (e.g., *B. subtilis* genes), for example, genes which have nucleic acids that are mutated, inserted or

30 deleted, but which encode proteins substantially similar to the naturally-occurring gene products of the present invention. For example, it is well understood that one of skill in the art can mutate (e.g., substitute) nucleic acids which, due to the degeneracy of the genetic code, encode for an identical amino acid as that encoded by the naturally-

35 occurring gene. Moreover, it is well understood that one of skill in the art can mutate (e.g., substitute) nucleic acids which encode for conservative amino acid substitutions. It is further well understood that one of skill in the art can substitute, add or delete



amino acids to a certain degree without substantially affecting the function of a gene product as compared with a naturally-occurring gene product, each instance of which is intended to be included within the scope of the present invention.

In a preferred embodiment, an isolated protein of the present invention (*e.g.*, an isolated pantothenate biosynthetic enzyme and/or an isolated isoleucine-valine biosynthetic enzyme and/or an isolated CoaA biosynthetic enzymes, for example isolated CoaA or CoaX) has an amino acid sequence shown in SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 or SEQ ID NO:87. In other embodiments, an isolated protein of the present invention is a homologue of the at least one of the proteins set forth as SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 or SEQ ID NO:87 (*e.g.*, comprises an amino acid sequence at least about 30-40% identical, preferably about 40-50% identical, more preferably about 50-60% identical, and even more preferably about 60-70%, 70-80%, 80-90%, 90-95% or more identical to the amino acid sequence of SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 or SEQ ID NO:87, and has an activity that is substantially similar to that of the protein encoded by the amino acid sequence of SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 or SEQ ID NO:87, respectively.

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical positions/total # of positions x 100), preferably taking into account the number of gaps and size of said gaps necessary to produce an optimal alignment.

The comparison of sequences and determination of percent homology between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-68, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-77. Such

an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Research* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller (1988) *Comput Appl Biosci.* 4:11-17. Such an algorithm is incorporated into the ALIGN program available, for example, at the GENESTREAM network server, IGH Montpellier, FRANCE (<http://vega.igh.cnrs.fr>) or at the ISREC server (<http://www.ch.embnet.org>). When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

In another preferred embodiment, the percent homology between two amino acid sequences can be determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 12, 10, 8, 6, or 4 and a length weight of 2, 3, or 4. In yet another preferred embodiment, the percent homology between two nucleic acid sequences can be accomplished using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a gap weight of 50 and a length weight of 3.

25

#### X. Biotransformations and Bioconversions

Another aspect of the present invention includes biotransformation processes which feature recombinant microorganisms (*e.g.*, mutant microorganisms) and/or isolated CoA, pantothenate or isoleucine-valine biosynthetic enzymes described herein. The term "biotransformation process", also referred to herein as "bioconversion processes", includes biological processes which result in the production (*e.g.*, transformation or conversion) of any compound (*e.g.*, intermediate or product) which is upstream of a CoA, pantothenate or isoleucine-valine biosynthetic enzyme to a compound (*e.g.*, substrate, intermediate or product) which is downstream of said CoA, pantothenate or isoleucine-valine biosynthetic enzyme.

35

In one embodiment, the invention features a biotransformation process for the production of a panto-compound comprising contacting a microorganism which overexpresses at least one pantothenate biosynthetic enzyme with at least one appropriate substrate or precursor under conditions such that said panto-compound is produced and recovering said panto-compound. In a preferred embodiment, the invention features a biotransformation process for the production of pantoate comprising contacting a microorganism which overexpresses ketopantoate reductase (the *panE* gene product) with an appropriate substrate (e.g., ketopantoate) under conditions such that pantoate is produced and recovering said pantoate. In another preferred embodiment, the invention features a biotransformation process for the production of pantothenate comprising contacting a microorganism which overexpresses ketopantoate reductase and pantothenate synthetase with appropriate substrates (e.g., ketopantoate and  $\beta$ -alanine) under conditions such that pantothenate is produced and recovering said pantothenate. In yet another preferred embodiment, the invention features a biotransformation process for the production of pantothenate comprising contacting a microorganism which overexpresses ketopantoate hydroxymethyltransferase, ketopantoate reductase and pantothenate synthetase with appropriate substrates (e.g.,  $\alpha$ -ketoisovalerate and  $\beta$ -alanine) under conditions such that pantothenate is produced and recovering said pantothenate. Preferred recombinant microorganisms for carrying out the above-described biotransformations include pantothenate kinase mutants. Conditions under which pantoate or pantothenate are produced can include any conditions which result in the desired production of pantoate or pantothenate, respectively.

In yet another embodiment, the present invention includes a method of producing  $\beta$ -alanine that includes culturing a microorganism which overexpresses aspartate- $\alpha$ -decarboxylase under conditions such that  $\beta$ -alanine is produced. Preferably, the aspartate- $\alpha$ -decarboxylase-overexpressing microorganism has a mutation in a nucleic acid sequence encoding a pantothenate biosynthetic enzyme selected from the group consisting of ketopantoate hydroxymethyltransferase, ketopantoate reductase and pantothenate synthetase.

The invention further features a method of producing  $\beta$ -alanine that includes contacting a composition comprising aspartate with an isolated *Bacillus* aspartate- $\alpha$ -decarboxylase enzyme under conditions such that  $\beta$ -alanine is produced (e.g., an *in vitro* synthesis method).

The microorganism(s) and/or enzymes used in the biotransformation reactions are in a form allowing them to perform their intended function (e.g., producing a desired compound). The microorganisms can be whole cells, or can be only those portions of the cells necessary to obtain the desired end result. The microorganisms can be

suspended (e.g., in an appropriate solution such as buffered solutions or media), rinsed (e.g., rinsed free of media from culturing the microorganism), acetone-dried, immobilized (e.g., with polyacrylamide gel or k-carrageenan or on synthetic supports, for example, beads, matrices and the like), fixed, cross-linked or permeablized (e.g., have permeablized membranes and/or walls such that compounds, for example, substrates, intermediates or products can more easily pass through said membrane or wall).

Purified or unpurified CoA biosynthetic enzyme(s) (e.g., CoaA and/or CoaX), pantothenate biosynthetic enzyme(s) and/or valine-isoleucine biosynthetic enzyme(s) can also be used in biotransformation reactions. The enzyme can be in a form that allows it to perform its intended function (e.g., obtaining the desired compound). For example, the enzyme can be in free form or immobilized. Purified or unpurified CoA biosynthetic enzyme(s), pantothenate biosynthetic enzyme(s) and/or valine-isoleucine biosynthetic enzyme(s) can be contacted in one or more *in vitro* reactions with appropriate substrate(s) such that the desired product is produced.

With respect to at least the above-described methodologies (e.g., the production methodologies of the present invention), at least one aspect of the invention features the following: embodiments in which the methods do not use microorganisms of the genus *Corynebacterium* and/or microorganisms of the genus *Escherichia*; embodiments in which the methods do not use microorganisms selected from the group consisting of *Escherichia coli* and *Corynebacterium glutamicum*; embodiments in which the methods do not use gram negative microorganisms; embodiments in which the microorganisms utilized do not include, express or produce nucleic acid molecules, genes or proteins (e.g., biosynthetic enzymes) derived from microorganisms of the genus *Corynebacterium* and/or microorganisms of the genus *Escherichia*; embodiments in which the microorganisms do not include, express or produce nucleic acid molecules, genes or proteins (e.g., biosynthetic enzymes) derived from microorganisms selected from the group consisting of *Escherichia coli* and *Corynebacterium glutamicum*.

#### XI. Screening Assays

Because CoA is an essential factor in bacteria, proteins (e.g., enzymes) involved in the biosynthesis of CoA provide valuable tools in the search for novel anti-biotics. In particular, the CoaX protein is a valuable target for identifying bacteriocidal compounds because it bears no resemblance in primary sequence to mammalian pantothenate kinase enzymes. Accordingly, the present invention also provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (e.g., peptides, peptidomimetics, small molecules or other drugs)

which bind to CoaX, or have a stimulatory or inhibitory effect on, for example, *coaX* expression or CoaX activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which are capable of binding to CoaX proteins or a biologically active  
5 portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which modulate the activity of CoaX proteins or biologically active portions thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or  
10 solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug*  
15 *Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem.*  
20 *Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233. Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci*  
25 *USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a microorganism-based assay in which a recombinant microorganism which expresses a CoaX protein or biologically active  
30 portion thereof is contacted with a test compound and the ability of the test compound to modulate CoaX activity is determined. Determining the ability of the test compound to modulate CoaX activity can be accomplished by monitoring, for example, intracellular phosphopantoate or CoA concentrations or secreted pantothenate concentrations (as compounds that inhibit CoaX will result in a buildup of pantothenate in the test  
35 microorganism). CoaX substrate can be labeled with a radioisotope or enzymatic label such that modulation of CoaX activity can be determined by detecting a conversion of labeled substrate to intermediate or product. For example, CoaX substrates can be

labeled with  $^{32}\text{P}$ ,  $^{14}\text{C}$ , or  $^3\text{H}$ , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Determining the ability of a compound to modulate CoaX activity can alternatively be determined by detecting the induction of a reporter gene (comprising a CoA-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*,  
5 luciferase), or detecting a CoA-regulated cellular response.

In yet another embodiment, a screening assay of the present invention is a cell-free assay in which a CoaX protein or biologically active portion thereof is contacted with a test compound *in vitro* and the ability of the test compound to bind to or modulate  
10 the activity of the CoaX protein or biologically active portion thereof is determined. In a preferred embodiment, the assay includes contacting the CoaX protein or biologically active portion thereof with known substrates to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to modulate enzymatic activity of the CoaX on its substrates.

15 Screening assays can be accomplished in any vessel suitable for containing the microorganisms, proteins, and/or reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either CoaX protein or a recombinant microorganism expressing CoaX protein to facilitate  
20 separation of products and/or substrates, as well as to accommodate automation of the assay. For example, glutathione-S-transferase/CoaX fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates. Other techniques for immobilizing proteins on matrices (*e.g.*, biotin-conjugation and streptavidin immobilization or antibody conjugation) can  
25 also be used in the screening assays of the invention.

In another embodiment, modulators of CoaX expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of *coaX* mRNA or CoaX polypeptide in the cell is determined. The level of expression in the presence of the candidate compound is compared to the level of expression in the  
30 absence of the candidate compound (or to a suitable control, for example, an appropriate buffer control or standard). The candidate compound can then be identified as a modulator of *coaX* mRNA or CoaX polypeptide expression based on this comparison.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an  
35 agent identified as described herein in an appropriate animal model. For example, an CoaX modulating agent identified as described herein (*e.g.*, an anti-bactericidal

compound) can be used in an infectious animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents, patent applications  
5 (including U.S. Patent Application Serial No. 09/400,494, filed September 21, 1999 (pending), provisional U.S. Patent Application Serial No. 60/210,072, filed June 7, 2000, provisional U.S. Patent Application Serial No. 60/221,938, filed July 28, 2000 and provisional U.S. Patent Application Serial No. 60/227,860, filed August 24, 2000, to which this application relates) and published patent applications cited throughout this  
10 application are incorporated herein by reference.

## EXAMPLES

### **General Methodology:**

**Strains.** *Bacillus subtilis* strains of the present invention are generally derived from either of two strains. The first is variously named "168", "1A1", or "RL-1". The genotype is *trpC2*. This strain was derived from the wild type "Marburg" strain by mutagenesis and has been the basis of much of the molecular biology work done on *B. subtilis*. The second strain is PY79, a prototrophic derivative of 168 that was made Trp<sup>+</sup> by transduction from the wild type strain W23.

**Media.** Standard minimal medium for *B. subtilis* is comprised of 1 x Spizizen salts and 0.5% glucose. Standard solid "rich medium" is Tryptone Blood Agar Broth (Difco), and standard liquid "rich medium" is VY, a mixture of veal infusion broth and yeast extract. For testing production of pantothenate in liquid test tube cultures, an enriched form of VY, called "Special VY" or "SVY" is used. For batch fermentations, SVYG and PFMG are used. The compositions of these media are given below.

**VY, a rich liquid medium:** 25 g Difco Veal Infusion Broth, 5 g Difco Yeast Extract, 1L water (autoclave).

**TBAB, a rich solid medium:** 33 g Difco Tryptone Blood Agar Broth, 1L water (autoclave).

**MIN, a minimal medium:** 100 ml 10 x Spizizen salts; 10 ml 50% glucose; 2 ml 10% arginine HCl\*; 10 ml 0.8% tryptophan\*\*; water to 1 liter. (\*In some cases, arginine is omitted or replaced by sodium glutamate at 0.04% final concentration. In general, *B. subtilis* grows faster in minimal medium when certain amino acids, such as arginine, glutamine, glutamate, or proline, are added as an auxiliary nitrogen source. ) (\*\*For strains that are tryptophan auxotrophs, tryptophan is routinely added to most minimal media.)

**10 x Spizizen Salts:** 174 g K<sub>2</sub>HPO<sub>4</sub>·3H<sub>2</sub>O; 20 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 60 g KH<sub>2</sub>PO<sub>4</sub>; 10 g Na<sub>3</sub>Citrate·2H<sub>2</sub>O; 2 g MgSO<sub>4</sub>·7H<sub>2</sub>O; water to 993 mls; then add 3.5 ml FeCl<sub>3</sub> solution and 3.5 ml Trace Elements solution.

**FeCl<sub>3</sub> Solution:** 4 g FeCl<sub>3</sub>·6H<sub>2</sub>O; 197 g Na<sub>3</sub>Citrate·2H<sub>2</sub>O; water to 1 liter (filter sterilize)

**Trace Elements Solution:** 0.15 g Na<sub>2</sub>MoO<sub>4</sub>·2H<sub>2</sub>O; 2.5 g H<sub>3</sub>BO<sub>3</sub>; 0.7 g CoCl<sub>2</sub>·6H<sub>2</sub>O; 0.25 g CuSO<sub>4</sub>·5H<sub>2</sub>O; 1.6 g MnCl<sub>2</sub>·4H<sub>2</sub>O; 0.3 g ZnSO<sub>4</sub>·7H<sub>2</sub>O; water to 1 liter (filter sterilize).



- SVY, Special VY, a supplemented\* rich medium for testing pantothenate production in test tube cultures:* 25 g Difco Veal Infusion Broth; 5 g Difco yeast extract; 5 g sodium glutamate; 2.7 g ammonium sulfate; 740 ml water (autoclave); add 200 ml 1 M potassium phosphate, pH 7.0; 60 ml 50% glucose. (\*For testing
- 5 pantothenate production in liquid SVY test tube cultures, Na  $\alpha$ -ketoisovalerate and/or  $\beta$ -alanine can be added to a concentration of 5 g/L from filter-sterilized stocks.)

- PFMG, a yeast extract based medium used in fermentors:* 20 g Amberex 1003™ yeast extract; 5 g sodium glutamate, 2 g ammonium sulfate; 5 g tryptophan; 10 g KH<sub>2</sub>PO<sub>4</sub>; 20 g K<sub>2</sub>HPO<sub>4</sub>·3H<sub>2</sub>O; 1 g MgCl<sub>2</sub>·6H<sub>2</sub>O; 0.1 g CaCl<sub>2</sub>·2H<sub>2</sub>O; 1 g sodium citrate;
- 10 0.01 g FeSO<sub>4</sub>·7H<sub>2</sub>O; 1 ml trace elements solution; 20 g glucose; add water to 1 L. Glucose or other sugars are fed as needed. Feed solutions can contain minerals, defined or food grade nutrients.

- PF, a chemically defined pantothenate free medium for testing pantothenate auxotrophy:* 100 ml 10 x Spizizen Salts; 100 ml 1 x Difco Pantothenate Assay Medium;
- 15 10 ml 50% glucose; water to 1 liter.

- For pantothenate auxotrophs, 1 mM Na pantothenate is added to both minimal and rich media, since there is generally not enough pantothenate in rich media to support *B. subtilis pan* mutants. Amino acids are at 100 mg per liter, when used. Selection for antibiotic resistance is done with 5 mg/L chloramphenicol, 100 mg/L
- 20 spectinomycin HCl, 15 mg/L tetracycline HCl, or 1 mg/L erythromycin plus 25 mg/L lincomycin.

- Pantothenate Assays: Biological assay.* The indicator organism, *Lactobacillus plantarum*, requires pantothenate for growth, and responds to low concentrations ( $\mu$ g/L).
- 25 Thus, using serial dilutions, a wide range of concentrations can be assayed. Commercially available medium (e.g., Pantothenate Assay Medium (PAM), Difco), can be used. However, Difco PAM supplemented with pantothenate does not support growth to the same level as obtainable using a fresh-mixed version of Pantothenate Assay Medium (FM-PAM), made up of the individual components as specified by
- 30 Difco, which is accordingly, routinely used instead of the commercial product.

- Before assaying *B. subtilis* culture supernatants, the *B. subtilis* cells must be either removed or killed. *B. subtilis* culture supernatants give approximately the same pantothenate titer when the supernatants are autoclaved as when they are sterile filtered. Accordingly, routine procedures involve autoclaving samples for 5 minutes prior to the
- 35 biological assay.

**Pantothenate Assays : HPLC assay.** Pantothenic acid production is measured by HPLC with a detector wavelength of 197 nm and a reference at 450 nm. The procedure is a modification of one recommended by Hewlett-Packard for water soluble vitamins. Samples of culture broth are diluted into an equal volume of 60% acetonitrile (ACN), centrifuged and filtered. Typically a further 10-fold dilution before analysis brings the final dilution to 20-fold. Higher concentrations of product are diluted further. Compounds are separated on a C18 Phenomenex 5 $\mu$  Aqua 250 x 4.6 mm column with 5% acetonitrile (ACN) in 50 mM Na phosphate buffer at pH 2.5. An ACN gradient from 5% to 95% washes the column between every sample. The area of the pantothenate peak is proportional to the concentration between 5 to 1000 mg/L. Other panto-compounds are also separated and quantitated by this method.

**Amino Acid Analysis: HPLC assay.** Amino acids present in the fermentation medium and throughout the fermentation are measured by HPLC with a detector wavelength of 338 nm and a reference at 390 nm. The procedure is a modification of one recommended by Hewlett-Packard for amino acid analysis. Samples of culture broth are prepared identically as for the panto-compound analysis. Compounds are separated on a C18 Hypersil 5 $\mu$  ODS 200 x 2.1 mm column. Solvent A is 20 mM Na acetate buffer at pH 7.2. Solvent B contains 40% ACN and 40% methanol. A gradient from 100% Solvent A to 100% Solvent B separates amino acids and washes the column between every sample.

**Batch Fermentations.** Pantothenate producing strains are grown in stirred tank fermentors, for example, in CF3000 Chemap 14 liter vessels with 10 liter working volumes. Computer control and data collection is by commercial software, for example, B. Braun Biotech MFCS software. Fermentations can be batch processes but are preferably sugar-limited, fed batch processes. Some media components (*e.g.* of SVYG and PMFG) are added to the fermentor and sterilized in place. Portions of the media are sterilized separately and added to the fermentors aseptically. This procedure is well known to those familiar with the art. Additional nitrogen sources in feeds are sterilized separately and added to the carbon source after cooling.

The initial sugar in the medium is consumed in approximately 6 hours. Afterwards, glucose or other sugars are fed with the possible addition of minerals, and defined or food grade nutrients. Alternatively, feeds are scheduled based on a consensus profile of nutritional requirements from samples taken from earlier fermentations.

After inoculation, agitation is set at a relatively low speed, e.g. 200 rpm. When the dissolved oxygen (pO<sub>2</sub>) falls to 30%, computer control automatically adjusts the agitation to maintain a dissolved oxygen concentration between 25 and 30% pO<sub>2</sub>.

**5 EXAMPLE I: Enhanced Production of a Panto-Compound Using Bacteria Overexpressing *panBCD* Gene Products.**

This Example describes the cloning of the *B. subtilis panBCD* operon and the generation of microorganisms overexpressing the *panBCD* gene products.

To clone the *B. subtilis panBCD* operon, a plasmid library of *B. subtilis* GP275  
10 (a derivative of 168) genomic DNA was transformed in *E. coli* BM4062 (*birA*<sup>ts</sup>), and temperature resistant clones were selected at 42°C. By comparison of restriction maps to the genome sequence, one particular clone was deduced to contain the *B. subtilis birA* gene and the adjacent *panBCD* genes. This plasmid was named pAN201.

To overexpress the *panBCD* operon and produce pantothenate, the native  
15 promoter of the *panBCD* operon was replaced by either of two strong, constitutive promoters derived from the *B. subtilis* bacteriophage SP01. These two promoters are named *P*<sub>26</sub> and *P*<sub>15</sub>. In addition, either of two artificial ribosome binding sites (RBSs) were used to replace the native *panB* RBS. These two artificial RBSs (set forth as SEQ ID NO:49 and SEQ ID NO:50) were predicted to increase translation of *panBCD*; their  
20 sequences are shown in Table 1A. Three such engineered *panBCD* expression cassettes were built into circular plasmids capable of replicating in *E. coli*. Other features of the plasmids include a strong rho-independent transcription terminator from the *E. coli* ribosomal RNA transcription unit, called T<sub>1</sub>T<sub>2</sub>, a Gram-positive chloramphenicol resistance gene (*cat*), derived from pC194, and a pair of *NotI* restriction sites at the  
25 junctions between the *E. coli* replicon and the segment intended for integration into *B. subtilis*. Three plasmids of this series, pAN004, pAN005, and pAN006 were constructed. pAN004 contains the *P*<sub>26</sub> promoter, RBS1, and a low copy *E. coli* replicon. pAN005 contains the *P*<sub>15</sub> promoter, which in our experience is not as strong as *P*<sub>26</sub>, RBS1, and the low copy replicon. pAN006 contains the *P*<sub>26</sub> promoter, RBS2, and a  
30 medium copy replicon.

The three *panBCD* expression cassettes contained in the above-mentioned three plasmids were all ligated to a DNA fragment consisting of sequences that naturally occur immediately upstream from the native *panB* gene and integrated in single copy by homologous recombination into the *panBCD* locus of *B. subtilis* strains RL-1 and PY79,  
35 replacing the wild-type operon. This was accomplished in two steps. First a deletion-substitution that replaced about two thirds of the *panB* coding region with a Gram-

positive spectinomycin resistance gene (*spec*) was integrated at *panB* to yield Spec<sup>r</sup>, pantothenate auxotrophs. These intermediate strains were then transformed with the *panBCD* expression cassettes of pAN004, pAN005, and pAN006 after ligating them to a DNA fragment containing chromosomal sequences just upstream of *panB*. Selection of the incoming cassette was for pantothenate prototrophy. The resulting strains were named PA221, PA222 and PA223 (from RL-1), and PA235, PA232 and PA233 (from PY79), respectively. An example of a plasmid that contains the joined upstream sequence that is in the integrated strain in PA221 is pAN240 (see Figure 2). The nucleotide sequence of pAN240 is set forth as SEQ ID NO:76.

10 Polymerase chain reaction using appropriate primers was used to verify the correct chromosomal structures of these engineered strains. When extracts of strain PA221 were examined by SDS-PAGE, two proteins were found to be overexpressed. One protein had an apparent molecular weight of 29,000 and the other protein appeared to be 39,000 daltons. The 29,000 dalton bands is presumably PanB (predicted molecular weight of 29,761). The larger protein band presumably represents PanC (predicted size 31,960 daltons).

20 The ability of these strains to produce pantothenate in test tube cultures was assessed as follows. Each strain was grown in SVY medium supplemented with 5 g/L  $\alpha$ -ketoisovalerate ( $\alpha$ -KIV) and 5 g/L  $\beta$ -alanine, to ensure that these precursors were not limiting. Culture supernatants were autoclaved and assayed using the bioassay. Relative to the parent strains, RL-1 and PY79, the engineered strains produced about 8- to 30-fold more pantothenate, attaining 1 g/L pantothenate in some cases.

Table 2. Production of pantothenate by engineered *B. subtilis* strains in liquid test tube cultures grown in SVY medium with 5 g/L  $\alpha$ -KIV and 5 g/L  $\beta$ -alanine.

Expt.	Strain	Promoter	RBS at <i>panB</i>	[pantothenate] mg/L
1	RL-1	Native	Native	30
	PA221	P <sub>26</sub>	RBS1	990
				790
	PA222	P <sub>15</sub>	RBS1	250
				250
	PA223	P <sub>26</sub>	RBS2	790
				790
2	PY79	Native	Native	40
	PA235	P <sub>26</sub>	RBS1	930
				860
	PA221	P <sub>26</sub>	RBS1	1100
				1030

- 5 The P<sub>26</sub> promoter was about 3- to 4-fold more effective than the P<sub>15</sub> promoter, while RBS1 and RBS2 were roughly equivalent. Plasmids such as pAN004, pAN005, pAN006 can also be recombined as circles into the *B. subtilis* wild type *panBCD* locus by Campbell-type (single crossover) integration, selecting for chloramphenicol resistance at 5 mg/L. Strains obtained in this fashion produce about the same amount of
- 10 pantothenate as strains PA221, PA222, and PA223, respectively. pAN004 containing the P<sub>26</sub> promoter, RBS1 and a low copy *E. coli* replicon, is depicted schematically in Figure 3A. The nucleotide sequence of plasmid pAN004 is set forth as SEQ ID NO:93. pAN006 containing the P<sub>26</sub> promoter, RBS2 and a medium copy *E. coli* replicon, is depicted schematically in Figure 3B. The nucleotide sequence of plasmid pAN006 is set
- 15 forth as SEQ ID NO:94. The nucleotide sequence of *panBCD* is set forth as SEQ ID NO:59 and the predicted amino acid sequences of PanB, PanC and PanD are set forth as SEQ ID NO:24, SEQ ID NO:26 and SEQ ID NO:28, respectively. Methods for manipulating *Bacilli* are described, for example, in Harwood, C.R. and Cutting, S.M. (editors), *Molecular Biological Methods for Bacillus* (1990) John Wiley & Sons, Ltd.,
- 20 Chichester, England, the content of which is incorporated herein by reference.

**EXAMPLE II: Enhanced Production of a Panto-Compound Using Bacteria Overexpressing the *panE1* Gene Product – Ketopantoate Reductase.**

This Example describes the cloning of the *B. subtilis panE1* gene and the generation of microorganisms overexpressing the *panE1* gene product.

- 5 Pan<sup>-</sup> *B. subtilis* strains (e.g., *B. subtilis* mutants blocked in the synthesis of pantothenic acid) had previously been isolated, one of which was reported to be affected in ketopantoate reductase activity (Baigori *et al.* (1991) *J. Bacteriol.* 173:4240-4242). However, the mutations in these strains were incorrectly mapped to the *purE-tre* interval of the *B. subtilis* genetic map which does not contain the *panE* or *panBCD* genes.
- 10 Furthermore as shown below, a *panE* mutant does not have a Pan<sup>-</sup> phenotype as the *ilvC* gene product can substitute for the *panE* gene product in *B. subtilis* as in other bacterial strains such as *E. coli*. More recently, the *S. typhimurium panE* gene has been located and determined to be allelic to *apbA*, a gene required for anaerobic purine biosynthesis (Frodyma *et al.* (1998) *J. Biol. Chem.* 273:5572-5576). *E. coli* carries a highly
- 15 homologous gene at the same map location. Identification of the *panE* genes in *E. coli* and *S. typhimurium* was complicated by the fact that the *ilvC* gene product, acetohydroxy acid isomeroreductase, is also capable of carrying out the ketopantoate reductase reaction. As a result, pantothenate auxotrophy is not obtained unless both *panE* and *ilvC* are mutated.
- 20 To identify the *B. subtilis panE1* gene, the *B. subtilis* genome was searched using the protein sequence of *E. coli* or *S. typhimurium* ApbA (PanE), and two open reading frames were identified having homology to ApbA, named *ylbQ* and *ykpB*. These genes were renamed *panE1* and *panE2*, due to their proposed function in pantothenate biosynthesis. Both *panE1* and *panE2* were cloned as PCR products generated from
- 25 RL-1 genomic DNA as a template. Both genes were disrupted by either a spectinomycin resistance gene (*spec*) or a chloramphenicol resistance gene (*cat*). The interrupted genes were each integrated by double crossover into PY79 to give PA240 ( $\Delta panE1::spec$ ) and PA241 ( $\Delta panE2::cat$ ). Neither of these strains were pantothenate auxotrophs when tested on pantothenate-free (PF) plates, although PA240 containing
- 30  $\Delta panE1::spec$  grew slightly more slowly on TBAB without added pantothenate than with a 1 mM pantothenate supplement. By comparison, a  $\Delta panB::spec$  strain does not produce single colonies on TBAB, presumably because *B. subtilis* has no active uptake system for pantothenate.

- 35 It was hypothesized that the *B. subtilis* gene, *ilvC*, could function for *panE* as had been shown for *E. coli*. Accordingly, the *panE1* and *panE2* disruptions were introduced into a strain, CU550, which is reported to be *trpC2 ilvC4 leuC124*. Both the single

*panE1* and the double *panE1*, *panE2* disruptants were pantothenate auxotrophs on PF medium.

5 Table 3. Phenotypes of various *panE1* and *panE2* mutants on rich and defined media.

Strain	Medium	Growth*:	
		- pan	+ pan
PY79	TBAB	+++	+++
	PF	++	++
PA240	TBAB spec	+	+++
	PF	++	++
PA241	TBAB cam	+++	+++
	PF	++	++
CU550	TBAB	+++	+++
	PF	++	++
PA256	TBAB spec	-	+++
	PF	-	++
PA258	TBAB spec, cam	-	+++
	PF	-	++

\*Each "+" represents about 1 mm of colony diameter after overnight at 37°C.

Thus, mutating both *panE1* and *ilvC* results in pantothenate auxotrophy, while mutating only *panE1* does not, similar to what has been reported for *E.coli* and *S. typhimurium*.

10 Next, the quantitative effect of *panE1* and *panE2* knockouts in a pantothenate overproducing strain (PA235 described herein) was examined. The *panE1* and *panE2* disruptions were introduced into PA235, either singly or together to produce PA245 ( $\Delta panE1::spec$ ), PA248 ( $\Delta panE2::cat$ ) and PA244 ( $\Delta panE1::cat$ ,  $\Delta panE2::spec$ ). The effect of each mutation on pantothenate production was then tested in liquid test tube cultures.

Table 4. *Pantothenate production by PA235 derivatives containing panE1 and panE2 disruptions.*

Strain	[pan] mg/L	% of PA235
PA235	990	(100)
PA235	940	95
PA245	59	6
PA245	82	8
PA248	1060	106
PA248	1030	104
PA244	25	3
PA244	50	5

Thus, deletion analysis indicated that the *panE1* gene contributes to over 90% of the pantothenate production, while deletion of *panE2* did not have a significant effect on pantothenate production. It is therefore concluded that *panE1* accounts for most, but not necessarily all, of the ketopantoate reductase activity in *B. subtilis*. The rest of the ketopantoate reductase activity is predicted to be supplied by *ilvC*.

Having identified *panE1* as an important gene for pantothenate production, increased *panE1* expression was tested to determine whether it could enhance pantothenate production in strains such as PA221 or PA235. The *panE1* coding sequence was installed downstream of the *P<sub>26</sub>* promoter and RBS2 in a vector, pOTP61, designed to integrate and amplify at either the *bpr* locus (a non-essential protease gene) or at the locus of the cloned insert. The resulting plasmid, pAN236 (Figure 4) was transformed into PA221, selecting for resistance to tetracycline at 15 mg/L. The nucleotide sequence of pAN236 is set forth as SEQ ID NO:77. One transformant, named PA236 was chosen for further study.

PA236 was shown to overexpress a protein of about 31,000 daltons, which is close to the expected molecular weight of 33,290 daltons for *panE1* protein. Briefly, whole cell extracts were prepared from PY79, RL-1, PA221, PA221/pOTP61 and PA236 (2 samples). Cell extracts were separated by gel electrophoresis and the gels were coomassie stained to visualize proteins. In cells engineered to overexpress *panE* (PA236-1 and PA236-2), a band was visible having an approximate molecular weight of ~31,000 daltons (as compared to molecular weight markers). Moreover, PA221 and PA236 expressed increased levels of a ~29,000 dalton band, corresponding to the *panB*



gene product, and a ~39,000 dalton band, presumably corresponding the *panC* gene product. Furthermore, *E. coli* transformed with pAN006 (Figure 3B) expressed bands correlating to the *panB* and *panC* gene products and *E. coli* transfected with PAN236 expressed a ~31,000 dalton band corresponding to the *panE* gene product.

- 5        Next, PA236 was compared to PA221 carrying the empty vector pOTP61 for pantothenate production in liquid test tube cultures supplemented with 5 g/L  $\beta$ -alanine and 5 g/L  $\alpha$ -KIV.

10        *Table 5.        Effect of overexpression of panE1 and panE2 on pantothenate production by engineered strains in liquid test tube cultures.*

Strain	Additional Plasmid	Gene Overexpressed	[Pantothenate] mg/L
PA221	pOTP61	none	1,000
			940
PA236	pAN236	<i>panE1</i>	2,030
			2,050
PA238	pAN238	<i>panE2</i>	530
			680

- Overexpression of *panE1* caused a two-fold increase in pantothenate production when compared to the parent strain (*e.g.*, to slightly over 2 g/L) whereas overexpression of *panE2* resulted in a strain that produced about 35% less pantothenate than the parent strain. The *panE1* nucleotide sequence and predicted amino acid sequence are set forth as SEQ ID NO:29 and SEQ ID NO:30.
- 15

**EXAMPLE III: Enhanced Production of a Panto-Compound by Culturing Bacteria Overexpressing *panE1* or *panBCD* in the Presence of Valine.**

- 20        The ability of valine to function as a media supplement (*e.g.*, as a substitute for  $\alpha$ -KIV) in strains engineered to overexpress the *panBCD* operon and *panE1* was evaluated. Valine is closely related to  $\alpha$ -KIV by transamination, is less expensive than  $\alpha$ -KIV, and is commercially available in kilogram quantities. Valine was substituted for  $\alpha$ -KIV in the standard liquid test tube cultures in SVY medium. The concentration of
- 25        valine was varied from 5 to 50 g/L. Although valine at 5 g/L was slightly less effective

than  $\alpha$ -KIV in promoting pantothenate production, valine at 10 or 20 g/L equaled or surpassed 5 g/L  $\alpha$ -KIV in promoting pantothenate production.

#### EXAMPLES IV-X Generation of Microorganisms Capable of Producing

##### 5 Pantothenate in a Precursor-Independent Manner

*B. subtilis* strains such as PA221 and PA235 (engineered to overexpress *panBCD*) and PA236 (engineered to overexpress *panBCD* and *panEI*) need to be fed  $\alpha$ -ketoisovalerate ( $\alpha$ -KIV) (or valine) and aspartate (or  $\beta$ -alanine) to achieve maximal pantothenate production, as both these precursors are limiting for pantothenate synthesis. Accordingly, manipulated microorganisms were designed to eliminate the need to feed limiting precursors of pantothenate biosynthesis in the production of pantothenate. These strains are also useful in the production of various pantothenate biosynthetic pathway intermediates.

##### 15 EXAMPLE IV: Generation of Microorganisms Capable of Producing Pantothenate in an Aspartate- (or $\beta$ -Alanine) Independent Manner

The *panD* gene was cloned into *B. subtilis* expression vector pOTP61 to construct pAN423 (Figure 5). The nucleotide sequence of pAN423 is set forth as SEQ ID NO:78. The *NotI* restriction fragment containing *panD* was isolated from pAN423, self ligated and used to transform PA221. Transformants resistant to Tet<sup>15</sup>, Tet<sup>30</sup>, and Tet<sup>60</sup> were isolated and saved for further analysis.

Six of the pAN423 transformants plus two control transformants were grown in SVY containing 5 g/l  $\alpha$ -KIV with and without 10 g/l aspartate and then assayed for pantothenate production (Table 6).

25

Table 6. Effect of overproducing PanD on pantothenate production with and without added aspartate.

Culture* (PA221 transformants)	Asp (10 g/L)	TetR** ( $\mu$ g/ml)	OD550	[pan] (mg/L)
pOTP61-1	-	60	8.0	76
pOTP61-2	-	60	7.7	91
423#1-1	-	15	8.5	180
423#1-2	-	15	8.0	150
423#1-3	-	30	8.3	220
423#1-4	-	30	8.5	280
423#1-5	-	60	8.9	580
423#1-6	-	60	8.8	280

- 70 -

pOTP61-1	+	60	7.5	380
pOTP61-2	+	60	6.9	560
423#1-1	+	15	8.5	1200
423#1-2	+	15	8.6	1000
423#1-3	+	30	8.8	1200
423#1-4	+	30	9.0	1200
423#1-5	+	60	9.0	1200
423#1-6	+	60	9.0	1200

\*Test tubes cultures were grown in SVY +  $\alpha$ -KIV (5 g/L) with Asp (10 g/L) where indicated.

\*\*TetR = Approximate Tet-resistance of transformant

The pAN423 transformants produced at least twice the amount of pantothenate as the controls (*i.e.*, to a level at or near that which was obtained in earlier experiments by the addition of  $\beta$ -alanine to the culture medium). The data also show that in the absence of added aspartate, transformants containing additional copies of the *panD* gene expression cassette produce more pantothenate than the control transformants. One of the transformants, 423#1-5, produced about five times as much pantothenate as the controls. These results indicated that increased levels of PanD protein "pull" the conversion of available aspartate towards  $\beta$ -alanine, and that increasing *panD* gene expression can result in enhancement of pantothenate production both in the presence and absence of added aspartate.

Transformant 423#1-5 was re-named strain PA401 and studied further in shake flask fermentations. The shake flask medium was SVY with maltose instead of SVY with glucose. Results of shake flask experiments agreed well with test tube experiments during the first 24 hours. In shake flask experiments without the addition of  $\beta$ -alanine, PA401 produced approximately 1.5 g/l of pantothenate in 24 hours. Addition of  $\beta$ -alanine to the culture medium did not further improve pantothenate titers (Table 7), indicating that with this strain and these fermentation conditions,  $\beta$ -alanine is not limiting pantothenate production. In fact, when no  $\beta$ -alanine is fed, one can observe that PA401 is secreting  $\beta$ -alanine in significant amounts into the medium.

Table 7. Shake flask cultures with strain PA401 (*panD*) with and without  $\beta$ -alanine.

Initial $\beta$ -ala Added	Amino acids (g/l)		24 hours		
	$\beta$ -ala	Val	pH	OD <sub>600</sub>	Pantothenate (g/l)
0	0.7	1.5	7.5	13.7	1.5
5 g/l	7.1	1.4	7.6	12.4	1.5

Each value represents the average of duplicate 250 ml baffled flasks containing 50 ml of medium, incubated at 37°C with shaking (200 rpm).

Base Medium: SVY with 10 g/l  $\alpha$ -KIV, 30 g/l maltose

2% Inoculum: SVY with Tet<sup>r</sup> grown 24 hours.

#### EXAMPLE V: Engineering the *panD* gene for Further Increased Synthesis of Aspartate Decarboxylase and Enhanced Production of Pantothenate

This Example describes the generation of improved ribosome binding sites (RBSs) in the *panD* gene to increase the translation of *panD* mRNA.

##### Increasing the translation of the *panD* gene mRNA by generation of synthetic *panD* RBSs

The RBS (SEQ ID NO:88) used to express *panD* in pAN423 is a synthetic RBS and has been used to successfully produce other proteins in *B. subtilis* at a high level. However, it contains six mismatches when aligned to the "ideal" *B. subtilis* RBS (SEQ ID NO:45) (e.g., an RBS having a sequence which is complementary to the 16S RNA sequence within the *B. subtilis* ribosome). (See e.g., Table 1B, mismatches in bold).

Two new RBSs were designed to more closely mimic the ideal RBS. These synthetic RBSs, named new design A (NDA) and new design B (NDB) (also referred to herein as RBS3 and RBS4), are set forth as SEQ ID NO:51 and SEQ ID NO:52 and are aligned with the ideal RBS in Table 1B.

Oligonucleotides corresponding to the top and bottom strands of each new RBS were synthesized, annealed, then used to replace the RBS in pAN420, generating plasmids pAN426 and pAN427. These constructions are illustrated in Figure 6. The presence of the NDA and NDB RBS in pAN426 and pAN427 was confirmed by DNA sequence analysis. Next, the *panD* genes from pAN426 and pAN427 were transferred to *B. subtilis* expression vector pOTP61 as shown in Figure 7, creating pAN428 and pAN429. The nucleotide sequence of pAN429 is set forth as SEQ ID NO:79.

*NotI* restriction fragments lacking the *E. coli* vector sequences were isolated from pAN428 and pAN429, self-ligated, and used to transform strain PA221 to resistance to Tet<sup>15</sup>. Four isolates resistant to Tet<sup>60</sup> were picked from each transformation and assayed for pantothenate and  $\beta$ -alanine production along with PA221 transformed with the empty vector (pOTP61) and PA221 transformed with pAN423 (strain PA401) (see Table 8).

Table 8. *Panthenate production by test tube cultures of PA221 transformed with pAN428 and pAN429*

Plasmid	Medium Supplements	OD550	Pan g/l	$\beta$ -Ala g/l
pOTP61	$\alpha$ -KIV <sup>5</sup>	10	UND	0.04
pAN423	$\alpha$ -KIV <sup>5</sup>	10	0.4	0.04
pAN428-1 *	$\alpha$ -KIV <sup>5</sup>	12	0.6	0.04
pAN428-2	$\alpha$ -KIV <sup>5</sup>	11	0.5	0.03
pAN428-3	$\alpha$ -KIV <sup>5</sup>	11	0.3	0.03
pAN428-4	$\alpha$ -KIV <sup>5</sup>	10	0.1	UND
pAN429-1	$\alpha$ -KIV <sup>5</sup>	12	0.6	0.04
pAN429-2	$\alpha$ -KIV <sup>5</sup>	11	0.5	0.04
pAN429-3	$\alpha$ -KIV <sup>5</sup>	11	0.6	0.05
pAN429-4 #	$\alpha$ -KIV <sup>5</sup>	12	0.8	0.10
pOTP61	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	11	0.5	0.08
pAN423	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	12	0.9	1.32
pAN428-1 *	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	12	0.8	1.97
pAN428-2	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	12	0.8	1.51
pAN428-3	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	12	0.9	1.02
pAN428-4	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	11	0.8	0.30
pAN429-1	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	12	0.8	1.78
pAN429-2	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	12	0.8	1.66
pAN429-3	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	12	0.8	1.78
pAN429-4 #	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	13	0.8	2.28

UND: Below the limits of detection. \* Renamed PA402 # Renamed PA403

When grown in medium supplemented with  $\alpha$ -KIV at 5 g/l ( $\alpha$ -KIV<sup>5</sup>), the pAN428-1 transformant and all four of the pAN429 transformants produced more

pantothenate than did PA401, suggesting that these transformants contain higher levels of aspartate decarboxylase activity. When grown in medium supplemented with  $\alpha$ -KIV<sup>5</sup> and Asp<sup>10</sup> none of the pAN428 or pAN429 transformants produced more pantothenate than PA401. However, the pAN428-1 transformant and all four of the pAN429 transformants produced significantly more  $\beta$ -alanine than did PA401. It is possible that the excess  $\beta$ -alanine produced from added aspartate causes inhibition of pantothenate production. Alternatively,  $\beta$ -alanine may accumulate because pantoate is limiting in these strains.

The strains that produced the highest level of  $\beta$ -alanine, the pAN428-1 and pAN429-4 transformants, were renamed PA402 and PA403, respectively. These two strains were grown in SVY medium supplemented with various intermediates and reassayed for pantothenate and  $\beta$ -alanine production. PA221 and PA401 were included as controls. The results of the assays are presented in Table 9.

Table 9. Pantothenate production of PA402 and PA403 in test tube cultures.

Strain	Medium Supplements	OD550	Pan g/l	$\beta$ -Ala g/l	Val g/l
PA221	$\alpha$ -KIV <sup>5</sup>	7.9	UND	UND	0.9
PA401	$\alpha$ -KIV <sup>5</sup>	8.7	0.3	0.04	0.9
PA402	$\alpha$ -KIV <sup>5</sup>	8.5	0.5	0.04	0.9
PA403	$\alpha$ -KIV <sup>5</sup>	9.4	0.7	0.07	0.9
PA221	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	9.8	0.4	0.11	0.8
PA401	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	9.1	0.8	1.15	0.8
PA402	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	9.4	0.8	2.02	0.8
PA403	$\alpha$ -KIV <sup>5</sup> + Asp <sup>10</sup>	9.7	0.7	2.40	0.8
PA221	Pantoate <sup>5</sup>	8.9	UND	UND	0.2
PA401	Pantoate <sup>5</sup>	8.7	0.3	0.02	0.2
PA402	Pantoate <sup>5</sup>	10.6	0.5	0.02	0.2
PA403	Pantoate <sup>5</sup>	10.5	0.7	0.02	0.2
PA221	Pantoate <sup>5</sup> + Asp <sup>10</sup>	9.5	0.4	0.06	0.2
PA401	Pantoate <sup>5</sup> + Asp <sup>10</sup>	9.2	2.2	0.62	0.2
PA402	Pantoate <sup>5</sup> + Asp <sup>10</sup>	9.1	2.8	1.17	0.2
PA403	Pantoate <sup>5</sup> + Asp <sup>10</sup>	10.2	2.9	1.58	0.2

UND: Below the limits of detection.

When grown in medium supplemented with either  $\alpha$ -KIV<sup>5</sup> or Pantoate<sup>5</sup>, PA402 and PA403 produced significantly more pantothenate than did PA401. As before, even though PA402 and PA403 produced significantly more  $\beta$ -alanine than PA401 when  
5 grown in medium supplemented with  $\alpha$ -KIV<sup>5</sup> and Asp<sup>10</sup>, they did not produce a proportional increase in pantothenate. However, when grown in medium supplemented with Pantoate<sup>5</sup> plus Asp<sup>10</sup>, both PA402 and PA403 produced significantly more pantothenate than PA401, about a 30% increase.

It can be concluded from these experiments that the improved NDA and NDB  
10 *panD* ribosome binding sites, engineered into pAN428 and pAN429, respectively, lead to increased levels of aspartate decarboxylase activity.

Increasing the translation of the *panD* gene mRNA by generation of synthetic *panD* RBSs within the *panBCD* operon

15 The native *B. subtilis panD* gene ribosome binding site (RBS) (SEQ ID NO:43), which is found in the *P*<sub>26</sub>*panBCD* operon cassette present in PA221 (and in other engineered pantothenate production strains described herein), is shown in Table 1C aligned with the ideal ribosome binding site (SEQ ID NO:47). The alignment shows  
20 mismatches between the native *B. subtilis panD* gene RBS, which is located within the coding sequence for PanC, and the the ideal RBS. Three new RBSs (within the P26 *panBCD* operon cassette) were generated to increase translation of the *panD* gene mRNA and to yield increased synthesis of aspartate decarboxylase. These synthetic RBSs (termed NDI, NDII, and NDIII, also referred to herein as RBS5, RBS6 and RBS7,  
25 respectively) are set forth as SEQ ID NO:55, SEQ ID NO:56 and SEQ ID NO:57, respectively) and are included in Table 1C. It should be noted that although changes in the *panD* RBS within the *panBCD* operon also changes the C-terminal amino acid sequence of the PanC protein encoded by that operon, an alignment of known and suspected PanC protein amino acid sequences showed that the sequence of the last nine  
30 amino acids of the *B. subtilis* PanC protein could be altered without affecting any conserved amino acid residues indicating that such changes should not reduce pantothenate synthetase activity or expression. The new RBSs were synthesized and incorporated into the *P*<sub>26</sub> *panBCD* operon expression cassette as follows.

First, PCR primers were designed to contain the following elements: (1) a  
35 nucleic acid sequence encoding the first five amino acids of PanD up to and including a unique *Bsi*WI restriction site that had been previously introduced into *panD* by PCR; (2)

a stop codon for *panC*, (3) at least one synthetic RBS; and (4) 30-39 bp of nucleic acid sequence having 100% identity with *panC* upstream of the *panD* RBS. The primers were named TP102, TP103, and TP104 and contain the NDI, NDII, and NDIII ribosome binding sites, respectively. These three primers were used in conjunction with the 5' primer TP101, which hybridizes near the start codon of *panC*, in three independent PCR reactions to generate the NDI, NDII, and NDIII PCR products. The PCR products were purified, digested with *Xba*I, then cloned into plasmid vector pASK-1BA3 which had been digested with *Xba*I and *Sma*I. The resulting plasmids were named pAN431, pAN432, and pAN433. The construction of pAN431 is illustrated in Figure 8 and is representative of all three plasmid constructions. The presence of the desired synthetic *panD* gene RBS in each new plasmid was confirmed by DNA sequencing.

Next, the modified *panC* genes containing the new *panD* RBSs were joined with the *panD* gene utilizing the unique *Bsi*WI restriction site. This was accomplished by isolating the appropriate *Nsi*I-*Bsi*WI restriction fragments from pAN431, pAN432, and pAN433 and ligating them with a 2395 bp *Nsi*I-*Bsi*WI restriction fragment from pAN420, which supplied the *Bsi*WI-modified *panD* gene. These constructions resulted in plasmids pAN441, pAN442, and pAN443, respectively. A representative construction (pAN441) is illustrated in Figure 9. The nucleotide sequence of pAN443 is set forth as SEQ ID NO:80.

The new *panD* gene RBSs were then substituted into the  $P_{26}$ *panBCD* operon expression cassette as follows. First, a deletion-insertion mutation which removes the region of *panC* containing the *panD* RBS was created. This was constructed by digesting pAN430 with a mixture of *Bsp*E1 and *Bgl*II and recovering the 4235 bp fragment which is now missing the 3' end of *panC* and the 5' end of *panD*. This fragment was ligated with an *Ava*I-*Bam*HI restriction fragment from plasmid pECC4, which contains the chloramphenicol acetyl transferase (*cat*) gene. The 5' extension produced by *Ava*I digestion is compatible with that produced by *Bsp*E1 while the *Bgl*II and *Bam*HI extensions are also compatible. The resulting plasmid was named pAN440, and its construction is illustrated in Figure 10.

The resulting deletion-insertion mutation was crossed into the  $P_{26}$ *panBCD* operon via homologous recombination by transforming PA221 with linearized pAN440 and selecting for resistance to chloramphenicol on Cam<sup>r</sup> plates containing 1 mM pantothenate. Several transformants were tested, and were all found to require 1 mM pantothenate for growth, as expected. Two of these transformants were named PA408A and PA408B and were assayed for pantothenate production. Neither strain synthesized measurable quantities of pantothenate, even when grown in medium



containing pantoate and  $\beta$ -alanine at 5 g/l, indicating that the strains are deficient in pantothenate synthetase activity. Next, the new *panD* RBSs were crossed into the *P*<sub>26</sub> *panBCD* operon by transforming PA408 with linearized pAN441, pAN442, and pAN443 plasmid DNA and selecting for growth on TBAB plates without pantothenate supplementation. A transformation with linearized pAN430 (including the native *panD* RBS) was included as a control and was expected to give rise to transformants identical to PA221 described herein. Four isolates from each transformation were assayed for pantothenate and  $\beta$ -alanine production in SVY medium supplemented with various intermediates (Tables 10 and 11).

Table 10. Pantothenate production of PA410 - PA413 in test tube cultures.

Strain	RBS	Medium Supplements	OD <sub>550</sub>	Pan g/l	$\beta$ -Ala g/l
PA221	native	Pantoate <sup>5</sup>	11	UND	UND
PA410-1	native	Pantoate <sup>5</sup>	12	UND	UND
PA410-2		Pantoate <sup>5</sup>	12	UND	UND
PA410-3		Pantoate <sup>5</sup>	12	UND	UND
PA410-4		Pantoate <sup>5</sup>	12	UND	UND
PA411-1	NDI	Pantoate <sup>5</sup>	12	0.23	UND
PA411-2		Pantoate <sup>5</sup>	12	0.20	UND
PA411-3		Pantoate <sup>5</sup>	12	0.19	UND
PA411-4		Pantoate <sup>5</sup>	12	UND	UND
PA412-1	NDII	Pantoate <sup>5</sup>	12	UND	UND
PA412-2		Pantoate <sup>5</sup>	11	UND	UND
PA412-3		Pantoate <sup>5</sup>	13	0.18	UND
PA412-4		Pantoate <sup>5</sup>	12	0.18	UND
PA413-1	NDIII	Pantoate <sup>5</sup>	12	0.18	UND
PA413-2		Pantoate <sup>5</sup>	12	0.17	UND
PA413-3		Pantoate <sup>5</sup>	12	0.16	UND
PA413-4		Pantoate <sup>5</sup>	12	0.17	UND

UND: Below the limits of detection.

Table 11. Pantothenate production of PA410 - PA413 in test tube cultures.

Strain	RBS	Medium Supplements	OD550	Pan g/l	$\beta$ -Ala g/l
PA221	native	Pantoate <sup>5</sup> + Asp <sup>10</sup>	11	0.3	UND
PA410-1	native	Pantoate <sup>5</sup> + Asp <sup>10</sup>	12	0.4	UND
PA410-2		Pantoate <sup>5</sup> + Asp <sup>10</sup>	12	0.4	UND
PA410-3		Pantoate <sup>5</sup> + Asp <sup>10</sup>	12	0.4	UND
PA410-4		Pantoate <sup>5</sup> + Asp <sup>10</sup>	12	0.4	UND
PA411-1	NDI	Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	1.7	0.4
PA411-2		Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	1.7	0.4
PA411-3		Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	1.8	0.3
PA411-4		Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	0.4	UND
PA412-1	NDII	Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	0.4	UND
PA412-2		Pantoate <sup>5</sup> + Asp <sup>10</sup>	12	0.4	UND
PA412-3		Pantoate <sup>5</sup> + Asp <sup>10</sup>	12	1.6	0.3
PA412-4		Pantoate <sup>5</sup> + Asp <sup>10</sup>	12	1.5	0.2
PA413-1	NDIII	Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	1.6	0.3
PA413-2		Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	1.6	0.4
PA413-3		Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	1.7	0.4
PA413-4		Pantoate <sup>5</sup> + Asp <sup>10</sup>	13	1.7	0.4

UND: Below the limits of detection.

- 5 As expected from previous experiments using PA221, none of the transformants that contained the native *panD* RBS produced measurable quantities of pantothenate when grown in medium supplemented with pantoate. However, nine of the twelve transformants expected to contain modified *panD* RBSs produced significant quantities of pantothenate (160-230 mg/l) under these conditions, indicating that they possess
- 10 elevated levels of aspartate decarboxylase activity. When grown in medium supplemented with both pantoate and aspartate, these same nine transformants produced approximately four times more pantothenate than those with the native *panD* RBS. In addition, these nine transformants accumulated measurable quantities of  $\beta$ -alanine (230-410 mg/l). All transformants produced roughly equivalent quantities of pantothenate
- 15 when grown in medium containing pantoate and  $\beta$ -alanine, demonstrating that each contains a functional pantothenate synthetase.

These data demonstrate that the synthetic *panD* RBSs are about four times more effective than the native *panD* RBS in directing translation of the *panD* gene mRNA and evidence the utility of such synthetic RBSs in enhancing pantothenate production. Additional approaches to increasing pantothenate production can include, for example, increasing the half-life of the *panD* gene mRNA, increasing the strength of the promoter for *panD* transcription and/or increasing the stability of the PanD protein.

**EXAMPLE VI: Construction of Strains Containing an Integrated  $P_{26}$  *panE1* Cassette without an Antibiotic Resistance Gene.**

10

Example II describes the identification of the *B. subtilis* *panE1* gene that encodes the enzyme responsible for the majority of the ketopantoate reductase activity in *B. subtilis*. PA236 (containing the pAN236 plasmid) produced about twice as much pantothenate (2 g/l) as its parent strain, PA221 (1 g/l) in 24 hour SVY test tube cultures. PA236 was presumed to contain an amplified (~3 copies) integrated pAN236 plasmid based on selection for tetracycline resistance (the *tetR* gene product being encoded on the pAN236 plasmid in addition to the  $P_{26}$  *panE1* cassette). Also useful in the methodologies of the present invention are strains that contain a single integrated unamplifiable copy of  $P_{26}$  *panE1* at the *panE1* locus, for example, without an antibiotic resistance gene in the strain. Such a strain was generated as follows.

20

A plasmid named pAN251 was derived from pAN236 by inserting additional chromosomal sequences just upstream and just downstream from the  $P_{26}$  *panE1* cassette. These additional sequences, which provide homology to allow integration of the  $P_{26}$  *panE1* cassette at *panE1* by double crossover, were obtained by PCR from chromosomal DNA as a template. pAN251 is shown in Figure 11. The nucleotide sequence of pAN251 is set forth as SEQ ID NO:81.

25

Next, a strain was constructed which allowed selection for the incoming  $P_{26}$  *panE1* cassette. The strain included the following three components: (1)  $P_{26}$  *panBCD*; (2)  $\Delta$ *panE1*; and (3) *ilvC*, since both *panE1* and *ilvC* must be mutated to have a Pan<sup>-</sup> phenotype. The starting strain was CU550 (*trpC2*, *ilvC4*, *leuC124*). The  $P_{26}$  *panBCD* cassette from PA221 chromosomal DNA was introduced in two steps to create strain PA290. Next,  $\Delta$ *panE1*::*spec* was transformed into PA290, using chromosomal DNA from strain PA240, to give strain PA294 (*trpC2*, *ilvC4*, *leuC124*,  $P_{26}$  *panBCD*,  $\Delta$ *panE1*::*spec*), which is a strict pantothenate auxotroph. Finally, PA294 was transformed with plasmid pAN251, selecting for pantothenate prototrophy, to give strain PA303. This strain was expected to have the genotype *trpC2*, *ilvC4*, *leuC124*,  $P_{26}$

30

35

*panBCD*,  $P_{26}$  *panE1*. PA303 was checked for the correct chromosomal structure at the *panE1* locus by PCR using primers that flank the  $P_{26}$  insertion just upstream of *panE1*. The PCR product from PA303 was of the expected size, with a concomitant loss of the PCR product from the wild type *panE1* gene, consistent with having obtained the  
 5 desired double crossover event. Furthermore, PA303 was tetracycline sensitive, which is also consistent with the desired double crossover event, as opposed to a Campbell-type single crossover of the plasmids into the chromosome. The *trp*, *ilv*, and *leu* auxotrophies from the parent strain were all maintained in PA303.

10 In 24 hour liquid SVY test tube cultures, PA303 produced almost the same level of pantothenate as positive control PA236, and about twice as much as PA221, which does not contain engineered *panE1* as indicated in Table 12.

Table 12. Pantothenate production by 24 hr. test tube cultures of PA303 and  
 15 controls grown in SVY plus 5 g/l  $\alpha$ -KIV and 5 g/l  $\beta$ -alanine.

Strain	OD <sub>600</sub>	[pan] g/l
PA221-1	10.9	0.85
PA221-2	10.5	0.85
PA236-1	9.5	1.74
PA236-2	9.3	1.70
PA303-1	10.8	1.66
PA303-2	10.7	1.61

#### EXAMPLE VII: Generation of Microorganisms Capable of Producing Pantothenate in an $\alpha$ -KIV (or Valine) Independent Manner

20  $\alpha$ -ketoisovalerate ( $\alpha$ -KIV) is a rate limiting intermediate for pantothenate production in certain strains deregulated for pantothenate synthesis. Addition of either  $\alpha$ -KIV or valine at 5 g/l increases pantothenate production about 5-fold in test tube cultures with strains such as PA221. In order to alleviate the need to feed either  $\alpha$ -KIV or valine, strains were engineered that have an increased capacity to synthesize  $\alpha$ -KIV.

25  $\alpha$ -KIV is produced in *B. subtilis* from pyruvate by the sequential action of three enzymes encoded by four genes, *ilvB* and *ilvN*, *ilvC*, and *ilvD*. In a wild type *B. subtilis*, three of the genes (*ilvB*, *ilvN*, and *ilvC*) are the first three genes of the large *ilv-leu* operon. The fourth gene necessary for  $\alpha$ -KIV synthesis, *ilvD*, is located by itself elsewhere on the chromosome. The *B. subtilis* *ilv-leu* operon is thought to be regulated

only by leucine levels. Feeding of exogenous leucine reduces transcription of the *ilv-leu* operon by about 13-fold, probably by an attenuation mechanism (Grandoni *et al.* (1992) *J. Bacteriol.* 174: 3212-3219). The only known feedback regulation in the *ilv-leu* pathway is the inhibition of the *leuA* gene product by leucine.

- 5       As a first step to deregulate the synthesis of  $\alpha$ -KIV, a copy of the *ilvBNC* region from the wild type *B. subtilis ilv-leu* operon was isolated by PCR, and installed adjacent to the  $P_{26}$  promoter and RBS2 on a vector, pOLL8, that was designed to integrate a single  $P_{26}$  expression cassette by double recombination at the *amyE* locus. The *amyE* gene encodes a nonessential  $\alpha$ -amylase, and is a useful locus for installing expression  
10 cassettes. The resulting plasmid, pAN267, is illustrated in Figure 12. The nucleotide sequence of pAN267 is set forth as SEQ ID NO:82. pAN267 readily gave stable transformants by double crossover at the *amyE* locus of *B. subtilis* strains, as described in detail below.

Construction of pantothenate overproducing strains that are leucine prototrophs

- 15       Initially, a *B. subtilis* strain containing *ilvC4* and  $\Delta$ *panE1* was used to introduce a single copy of  $P_{26}$  *panE1* into the chromosome without using an antibiotic resistance gene. The double mutant was required to select for the incoming  $P_{26}$  *panE1* cassette because a  $\Delta$ *panE1* mutation alone does not result in pantothenate auxotrophy. A strain named CU550 was obtained containing *ilvC4* to be used as a basis for this type of strain  
20 construction. However, CU550 also contains a closely linked *leuC124* mutation, so all strains derived from CU550 required leucine. Having shown that the combination of  $P_{26}$  *panBCD* and  $P_{26}$  *panE1* was favorable for pantothenate production, the next step was to reassemble this combination of two cassettes in a leucine prototroph.

- Accordingly, the two cassettes were combined in two different strain  
25 backgrounds, RL-1 and PY79. To introduce chromosomal  $P_{26}$  *panE1* into the PY79 and RL-1 strain backgrounds without using an antibiotic resistance gene, a strategy was used that did not rely on *ilvC4*. (The strategy took advantage of the observation that the  $\Delta$ *panE1* mutation causes a pantothenate bradytroph, manifested by relatively small colonies on TBAB (rich) plates). First,  $\Delta$ *panB::cat* and  $\Delta$ *panE::spec* were introduced  
30 into both strain backgrounds. Next, the resulting strains were transformed simultaneously with DNA from two strains, PA221 ( $P_{26}$  *panBCD*) and PA303 ( $P_{26}$  *panE1*), selecting for Pan<sup>+</sup> on TBAB plates. Colonies of two distinct sizes grew on the selective plates, with the larger size comprising about 2% of the colonies. The larger colonies were presumed to represented co-transformants that received both  $P_{26}$  *panBCD*  
35 and  $P_{26}$  *panE1*, and that the smaller colonies had received only  $P_{26}$  *panBCD*. Consistent

with this prediction, the larger colonies had lost both *Cam*<sup>r</sup> and *Spec*<sup>r</sup>, while the smaller colonies had lost only the *cat* gene, and retained the *spec* gene. Furthermore, a representative derivative of PY79 named PA327, and a representative derivative of RL-1, named PA328, both produced the elevated levels of pantothenate in test tube cultures which was about 1.6 to 1.7 g/l (Table 13).

Table 13. *Pantothenate production of PA327, PA328, and controls from 24 hr test tube cultures grown in SVY plus 5 g/l  $\alpha$ -KIV and  $\beta$ -alanine.*

Strain	Background	<i>P26 panE1</i> copy number	[pan] g/l
PA221-1	RL-1	0	0.92
PA221-2	RL-1	0	0.95
PA236-1	RL-1	amplified (~3)	1.60
PA236-2	RL-1	amplified (~3)	1.73
PA327-1	PY79	1	1.66
PA327-2	PY79	1	1.65
PA328-1	RL-1	1	1.61
PA328-2	RL-1	1	1.91

10

Thus, PA327 and PA328 were concluded to contain both *P26 panBCD* and *P26 panE1*, and were used for further constructions as described below. PCR analysis confirmed the presence of the two cassettes.

15 Installation of a stable *P26 ilvBNC* cassette into two lineages of pantothenate overproducing strains

Having constructed PA327 and PA328, derivatives of PY79 and RL-1 that contain *P26 panBCD* and *P26 panE1*, and that are *Leu*<sup>+</sup>, the next step was to introduce stable copies of *P26 ilvBNC*. This was accomplished by transforming PA327 and PA328 with plasmid pAN267, selecting for *Spec*<sup>r</sup>. Screening by PCR showed that about 85% of the obtained transformants contain *P26 ilvBNC* integrated at *amyE* by double crossover. One transformant of PA327, named PA340, and one transformant of PA328, named PA342, were chosen for further study.

25 In test tube cultures grown in SVY medium plus 5 g/l  $\beta$ -alanine but without added  $\alpha$ -KIV, both PA340 and PA342 gave the expected increase in pantothenate production over that of PA327 and PA328, to about 1.3 to 2 g/l (Table 14).

Table 14. *Pantothenate and valine production by PA340 and PA342, both containing P<sub>26</sub> ilvBNC in 24 hr test tube cultures grown in SVY with 5 g/l  $\beta$ -alanine and with or without 5 g/l  $\alpha$ -KIV*

5

Strain	Back-ground	OD <sub>600</sub>		[pan] g/l		[val] g/l	
		- $\alpha$ -KIV	+ $\alpha$ -KIV	- $\alpha$ -KIV	+ $\alpha$ -KIV	- $\alpha$ -KIV	+ $\alpha$ -KIV
PA340-1	PY79	11.8	7.1	2.02	2.10	0.38	0.90
PA340-2	PY79	10.3	7.5	1.97	2.03	0.40	0.91
PA342-1	RL-1	10.2	8.0	1.29	1.89	0.27	0.78
PA342-2	RL-1	9.6	9.2	1.34	2.04	0.21	0.79

The two new strains also gave a slight increase in valine secretion, indicating that the *ilvBNC* genes had been deregulated. However, when the same strains were grown with 5 g/l  $\alpha$ -KIV added, a further increase in pantothenate production occurred from PA342, suggesting that  $\alpha$ -KIV was still rate limiting in this strain background. Similar results, only with more growth and hence higher pantothenate levels, were seen in shake flask cultures (Table 15).

Table 15. *Pantothenate and valine production by PA340 and PA342, both containing P<sub>26</sub> ilvBNC in 24 hour shake flask cultures grown in SVY with 5 g/l  $\beta$ -alanine and with or without 5 g/l  $\alpha$ -KIV.*

15

Strain	Back-ground	OD <sub>600</sub>		[pan] g/l		[val] g/l	
		- $\alpha$ -KIV	+ $\alpha$ -KIV	- $\alpha$ -KIV	+ $\alpha$ -KIV	- $\alpha$ -KIV	+ $\alpha$ -KIV
PA327	PY79	21	22	0.6	3.0	0.5	1.3
PA340-1	PY79	20	20	3.5	4.1	1.0	1.9
PA340-2	PY79	22	19	3.0	2.1	0.8	1.4
PA328	RL-1	20	16	1.4	2.7	0.6	1.3
PA342-1	RL-1	17	16	3.3	3.6	0.9	1.6
PA342-2	RL-1	18	18	3.1	4.2	0.8	1.4

**EXAMPLE VIII: Increasing *panD* Copy Number in Strains Engineered to Overproduce *panE1* and the *ilvBNC* Gene Products Enhances Pantothenate Production**

- Experiments where  $\beta$ -alanine was fed to cultures of engineered *B. subtilis* strains consistently showed that  $\beta$ -alanine was a rate limiting intermediate in pantothenate synthesis. The effect of adding additional copies of *panD* on pantothenate production in PA340 and PA342 was examined. Strains PA340 and PA342 were transformed with chromosomal DNA isolated from PA401 with selection on plates containing 15  $\mu$ g/ml of tetracycline (Tet<sup>15</sup> plates). Transformants derived from each parent were patched onto Tet<sup>60</sup> plates to identify those which were likely to contain multiple copies of the expression cassette. Twelve transformants from each transformation which grew on Tet<sup>60</sup> were streaked for single colonies on this medium and then assayed in SVY medium test tube cultures for pantothenate production. One transformant from each group was found to produce greater than 300 mg/l pantothenate in 24 hours. These two transformants were saved and named PA404 (PA340 strain background) and PA405 (PA342 strain background). Both strains were resistant to spectinomycin, indicating that the *P<sub>26</sub> ilvBNC* expression cassette was still present at *amyE*. PCR analysis of chromosomal DNA isolated from each strain confirmed that the deregulated *panE1* gene had also been retained.
- Next, PA404 and PA405 were evaluated in shake flask cultures which were grown in SVY medium containing maltose as the carbon source and supplemented with various intermediates. The cultures were grown for 24 and 48 hours and then assayed for pantothenate,  $\beta$ -alanine, and valine production. The results of this experiment are presented in Table 16. Analogous shake flask culture data for the parent strains (PA340 and PA342) are included in the tables for comparison.

Table 16. Pantothenate production by PA404 and PA405 in shake flask cultures after 24 hours

Strain	Medium Supplements	OD <sub>600</sub>	Pan g/l	$\beta$ -Ala g/l	Val g/l
PA340	none	20	0.4	<0.1	1.0
PA404	none	22	1.8	<0.1	0.7
PA342	none	19	0.3	0.2	0.7
PA405	none	19	1.4	0.4	0.5



- 84 -

PA340	$\beta$ -alanine <sup>5</sup>	18	3.6	3.2	0.6
PA404	$\beta$ -alanine <sup>5</sup>	18	2.8	5.1	0.7
PA342*	$\beta$ -alanine <sup>5</sup>	17	3.3	3.3	0.5
PA405*	$\beta$ -alanine <sup>5</sup>	19	1.3	6.5	0.6

Values are the average of duplicate flasks except where indicated by \*.

In the absence of any medium supplementation, PA404 and PA405 made four to five times more pantothenate in 24 hours compared to their isogenic parent strains (Table 16). The supply of  $\beta$ -alanine was clearly limiting in the parent strains PA340 and PA342. Addition of amplified *P26 panD* greatly increased the supply of  $\beta$ -alanine.

#### EXAMPLE IX: Deregulation of the *B. subtilis ilvD* Gene Enhances Pantothenate Production

To deregulate expression of the *ilvD* gene, standard procedures (described above) were used to integrate the constitutive *P26* promoter and an artificial ribosome binding site, RBS2, just upstream of the *ilvD* coding region. The *ilvD* gene maps by itself, unlinked to the *ilvBNC* operon. First, a 2.4 kb region of the RL-1 chromosome that contains the *ilvD* coding region and 730 bp of upstream sequence was cloned by PCR into a low copy (about 15 per *E. coli* cell) vector called pOK12, to give plasmid pAN257, shown in Figure 13.

Taking advantage of a natural *EcoRI* site just upstream of the native *ilvD* gene promoter, and a natural *NcoI* site at the *ilvD* start codon, an artificial sequence containing *P26* and RBS2 was inserted into pAN257 to give pAN263 (Figure 14). The nucleotide sequence of pAN263 is set forth as SEQ ID NO:83. In parallel with this construction, the *cat* gene was also inserted into pAN257, between the same upstream *EcoRI* site and a *BglII* site in the middle of the *ilvD* coding region, to give pAN261, which is deleted for a large portion of the *ilvD* gene (Figure 15).

Using pAN261 and pAN263, the *P26 ilvD* cassette could then be installed in the *B. subtilis* chromosome in two steps. In the first step, pAN261 is introduced by transformation, selecting for chloramphenicol resistance, and then confirming an *Ilv*<sup>-</sup> phenotype. In the second step, pAN263 is introduced, selecting for *Ilv*<sup>+</sup>, checking for chloramphenicol sensitivity, and confirming correct local structure by PCR.

pAN261 was first transformed into strain RL-1 (highly competent) to give strain PA343 ( $\Delta ilvD::cat$ ), and then chromosomal DNA from PA343 was used to transform PA340 and PA342 to *Ilv*<sup>-</sup> auxotrophy, yielding strains named PA348 and PA349, respectively. Chromosomal DNA is inherently more efficient than monomeric plasmid

in transforming *B. subtilis*. Similarly, pAN263 DNA was transformed into PA343 (moderately competent) to give strain PA345 ( $P_{26}$  *ilvD*), and then PA345 chromosomal DNA was used to transform PA348 and PA349 to *Ilv*<sup>+</sup> prototrophy, yielding strains PA374 and PA354, respectively.

- 5 As predicted, PA374 and PA354 gave further increases in pantothenate production, to about 2.5 to 2.9 g/l, in test tube cultures grown in SVY plus 5 g/l  $\beta$ -alanine (Table 17).

10 Table 17. Pantothenate and valine production by PA374 and PA354, containing  $P_{26}$  *ilvD*, and controls, in 24 hr test tube cultures grown in SVY with 5 g/l  $\beta$ -alanine and with or without 5 g/l  $\alpha$ -KIV.

Strain	Back-ground	<i>ilvD</i> status	OD <sub>600</sub>		[pan] g/l		[val] g/l	
			$\alpha$ -KIV -	+	$\alpha$ -KIV -	+	$\alpha$ -KIV -	+
PA340	PY79	w.t.	9.2	9.0	2.14	2.23	0.38	0.90
PA348	PY79	<i>ilvD::cat</i>	11.7	10.0	0.19	2.23	0.19	0.91
PA374-1	PY79	$P_{26}$ <i>ilvD</i>	9.1	7.3	2.93	2.40	0.58	0.87
PA374-2	PY79	$P_{26}$ <i>ilvD</i>	8.2	7.7	2.99	2.36	0.60	0.95
PA342	RL-1	w.t.	10.2	8.0	1.29	1.89	0.27	0.78
PA349	RL-1	<i>ilvD::cat</i>	8.1	7.7	0.17	1.87	0.22	0.88
PA354-1	RL-1	$P_{26}$ <i>ilvD</i>	9.6	9.6	2.57	2.03	0.65	1.23
PA354-2	RL-1	$P_{26}$ <i>ilvD</i>	7.5	8.2	2.48	2.24	0.64	0.97

- 15 In the absence of added  $\beta$ -alanine, strains PA374 and PA354 produced only about 0.2 g/l pantothenate in test tube cultures, indicating that PanD activity is significantly rate limiting.

- 20 To alleviate this limitation, the amplifiable  $P_{26}$  *panD* cassette from strain PA401 was installed. PA401 chromosomal DNA was transformed into PA374 and PA354, selecting for Tet<sup>r</sup> at 15 mg/l, to yield strains PA377 and PA365, respectively. After transformants were obtained, the strains were streaked on plates containing 30 and 60 mg/l tetracycline to reamplify the copy number of the  $P_{26}$  *panD* cassette integrated at the *bpr* locus. In test tube cultures grown in SVY without  $\alpha$ -KIV or  $\beta$ -alanine, a substantial improvement in pantothenate titers over those of PA374 and PA354 was obtained (Tables 18 and 19).

Table 18. *Pantothenate production by PA365, containing amplified P<sub>26</sub> panD, and controls, in 24 and 36 hr test tube cultures grown in SVY-glucose without  $\beta$ -alanine or  $\alpha$ -KIV.*

Strain	Relevant genotype	OD <sub>600</sub>		[pan] g/l	
		24 hrs.	36 hrs	24 hrs.	36 hrs.
PA342-1-1	w.t. <i>ilvD</i>	11.7	8.8	b.d.	0.27
PA342-1-2	w.t. <i>ilvD</i>	12.8	8.8	b.d.	0.26
PA354-1-1	P <sub>26</sub> <i>ilvD</i>	n.d.	11.0	n.d.	0.19
PA354-1-2	P <sub>26</sub> <i>ilvD</i>	n.d.	8.4	n.d.	0.20
PA365-1	P <sub>26</sub> <i>ilvD</i> , P <sub>26</sub> <i>panD</i>	9.8	10.0	1.01	2.07
PA365-2	P <sub>26</sub> <i>ilvD</i> , P <sub>26</sub> <i>panD</i>	9.9	10.4	0.96	2.09

5

n.d. = not determined; b.d. = below detection

Table 19. *Pantothenate production by PA377, containing amplified P<sub>26</sub> panD, and controls, in 27 hr test tube cultures grown in SVY-glucose or SVY-maltose, without  $\alpha$ -KIV, and with or without  $\beta$ -alanine.*

10

Strain	Relevant genotype	OD <sub>600</sub>			
		- $\beta$ -ala Glucose	+ $\beta$ -ala Glucose	- $\beta$ -ala Maltose	+ $\beta$ -ala Maltose
PA374-1	P <sub>26</sub> <i>ilvD</i>	9.4	9.8	7.0	6.4
PA374-2	P <sub>26</sub> <i>ilvD</i>	9.2	9.6	6.6	6.3
PA377-1	P <sub>26</sub> <i>ilvD</i> , P <sub>26</sub> <i>panD</i>	10.0	7.6	7.2	6.1
PA377-2	P <sub>26</sub> <i>ilvD</i> , P <sub>26</sub> <i>panD</i>	10.5	7.8	9.4	5.4

Strain	Relevant genotype	[pan] g/l			
		- $\beta$ -ala Glucose	+ $\beta$ -ala Glucose	- $\beta$ -ala Maltose	+ $\beta$ -ala Maltose
PA374-1	P <sub>26</sub> <i>ilvD</i>	0.04	2.76	0.14	1.31
PA374-2	P <sub>26</sub> <i>ilvD</i>	0.10	2.65	0.15	1.33
PA377-1	P <sub>26</sub> <i>ilvD</i> , P <sub>26</sub> <i>panD</i>	1.25	2.76	1.26	1.10
PA377-2	P <sub>26</sub> <i>ilvD</i> , P <sub>26</sub> <i>panD</i>	1.25	2.35	1.31	1.26

15 In SVY with glucose, an increase in pantothenate production can still be achieved by feeding 5 g/l  $\beta$ -alanine suggesting that increasing *panD* expression further might increase pantothenate production. In SVY with maltose, no further increase in pantothenate was obtained by feeding  $\beta$ -alanine suggesting that  $\beta$ -alanine and/or

aspartate synthesis is suppressed by glucose. Strains PA377 and PA365 have been evaluated in 10 liter fermentors, where they typically produce above 20 g/l pantothenate in 48 hours without supplemental  $\beta$ -alanine and  $\alpha$ -KIV or valine, described in detail below.

5

#### EXAMPLE X: 10 liter Fermentations of Pantothenate-Producing Microbes

Engineering of the  $P_{26}$  *ilvBNC* and  $P_{26}$  *ilvD* cassettes to give strains PA342 and PA354 allowed the production of 22 and 26 g/l of pantothenate, respectively, without the addition of valine or  $\alpha$ -KIV to the fermentation medium (Table 20). At 48 hours,

10 both strains had secreted about 0.5 g/l of valine into the medium.

Table 20. 10-liter fermentations of five pantothenate overproducing strains.

Strain	Medium	Feed 40% Glucose plus	OD 600 48 hr	Valine 48 hours g/l	$\beta$ -ala 48 hr g/l	Pantothenate g/L		
						36 hr	48 hr	72 hr
PA 236	SVYG	50 g/l $\beta$ -ala 25 g/l $\alpha$ -KIV	108	added	added	16	19	21
PA 342	SVYG	50 g/l $\beta$ -ala	92	0.5	added	17	22	--
PA 354	SVYG	50 g/l $\beta$ -ala	90	0.5	added	19	26	--
PA 365	SVYG	25g/l YE	77	0.85	0.4	18	21	27
PA 377	SVYG	25g/l YE	85	1.5	0.5	18	22	31
PA 377	PFMG	25g/l YE	96	0.8	0.4	19	25	29
PA377	PFMG	-	71	0.7	0.1	16	21	-

#### 15 Pantothenate synthesis in fermentors

With the addition of the  $P_{26}$  *panD* cassette to strains PA354 and PA374 to create strains PA365 and PA377, neither  $\beta$ -alanine nor  $\alpha$ -KIV needed to be added to the fermentors. Strain PA365 produced 21 g/l pantothenate in 48 hours and 27 g/l in 72 hours with no precursors added to the medium (Table 20). PA377 was somewhat better,

20 producing 18 g/l of pantothenate in 36 hours, 22 g/l in 48 hours, and 31 g/l in 72 hours). Valine was measured at 0.85 and 1.5 g/l for strains PA365 and PA377, respectively, at

48 hours in SVYG medium. Strain PA377 maintained valine between 1-1.5 g/l throughout most of the fermentation and  $\beta$ -alanine between 0.2 and 0.5 g/l.

Strain PA377 was further evaluated in 10-liter fermentors in yeast extract based PFMG medium. Pantothenate yields in PFMG and SVYG medium were similar. In PFMG, PA377 produced 19 g/l of pantothenate in 36 hours, 25 g/l in 48 hours, and 29 g/l in 72 hours. In SVYG, PA377 produced 18 g/L pantothenate in 36 hours, 22 g/L in 48 hours and 31 g/L in 72 hours (Table 20).

#### EXAMPLE XI: Converting Strain PA377 to a Tryptophan Prototroph

PA377 (Trp<sup>-</sup>) was transformed to Trp<sup>+</sup> using chromosomal DNA from PY79 to give strain PA824. After re-amplification of the *P<sub>26</sub>panD* cassette, PA824 was compared to PA377 for pantothenate production in test tube cultures grown in SVY glucose with or without 5 g/L  $\beta$ -alanine (Table 21).

Table 21 : *Trp<sup>+</sup> derivatives of PA377: Pantothenate production in 48 hour test tube cultures grown in SVY glucose,  $\pm\beta$ -alanine*

Strain	<i>trpC</i> donor	OD <sub>600</sub>		[pan] g/L	
		- $\beta$ -alanine	+ $\beta$ -alanine	- $\beta$ -alanine	+ $\beta$ -alanine
PA377-1	RL-1	8	8	1.5	3.4
PA377-2	RL-1	8	9	1.6	3.6
PA824-1	PY79	12	10	0.7	3.7
PA824-2	PY79	11	11	1.9	4.9

The Trp<sup>+</sup> strains grew to slightly higher densities than PA377. In the absence of exogenous  $\beta$ -alanine, all of the strains produced similar levels of pantothenate, while with the addition of  $\beta$ -alanine, the Trp<sup>+</sup> derivatives produced somewhat more pantothenate.

#### 25 Fermentor studies with PA824

PA824 was evaluated in CF3000 Chemap 14 liter vessels with 10 liter working volumes. Formulations for two of the media used in the fermentors are given in Tables 22 and 23.

Table 22 : Formulation for PFMG-5 medium

BATCH		
	MATERIAL	g/L (final [I])
1	Amberex 1003	10
2	Na Glutamate	5
3	$(\text{NH}_4)_2\text{SO}_4$	8
4	MAZU DF 37C	2.5
Added After Sterilization and Cool Down		
1	$\text{KH}_2\text{PO}_4$	10
2	$\text{K}_2\text{HPO}_4 \cdot 3\text{H}_2\text{O}$	20
1	Glucose	20
2	$\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$	1
3	$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	0.1
1	Sodium Citrate	1
2	$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	0.01
3	SM-1000X	1.0 ml
	$\text{H}_2\text{O}$	qs to 6000 ml

5

FEED		
	MATERIAL	g/L
1	Glucose	600
2	$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	0.6
	$\text{H}_2\text{O}$	qs to 3000 ml

Table 23 : Formulation for SVY-4 medium

BATCH		
	MATERIAL	g/L (final [])
1	Veal Infusion	25
2	Yeast Extract	5
3	Na Glutamate	5
4	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	4
5	MAZU DF 37C	2.5

## Added After Sterilization and Cool Down

1	KH <sub>2</sub> PO <sub>4</sub>	10
2	K <sub>2</sub> HPO <sub>4</sub> ·3H <sub>2</sub> O	20
1	Glucose	20
2	MgCl <sub>2</sub> ·6H <sub>2</sub> O	1
3	CaCl <sub>2</sub> ·2H <sub>2</sub> O	0.1
1	Sodium Citrate	1
2	FeSO <sub>4</sub> ·7H <sub>2</sub> O	0.01
3	SM-1000X	1.0 ml
	H <sub>2</sub> O	qs to 6000 ml

5

## FEED

	MATERIAL	g/L
1	Glucose	600
2	CaCl <sub>2</sub> ·2H <sub>2</sub> O	0.6
	H <sub>2</sub> O	qs to 3000 ml

All fermentations were glucose limited fed batch processes. Immediately after inoculation, agitation was set at 200 rpm. The initial batched 2% glucose was consumed during exponential growth. Afterwards, glucose concentrations were maintained

10

between 0.2 and 1.0 g/L by continuous feeding of a 60% glucose solution. The variable rate feed pump was computer controlled and linked to the dissolved oxygen concentration  $[pO_2]$  in the tank by an algorithm. When the  $[pO_2]$  fell to 30%, computer control began to automatically adjust the agitation rate to maintain a dissolved oxygen concentration between 25 and 30%  $[pO_2]$ . Computer control and data recording were by Braun MFCS software.

In one study, PA284 was grown in fermentors at two temperatures (40°C and 43°C) in the medium described in Table 22. Results of two experiments demonstrated that the highest pantothenate titers at early time points were produced at 43°C. The cell mass approached 150 optical density units at  $OD_{600}$  and 56 hours at 43°C, and the pantothenate titers were 21 g/L, 28 g/L and 36 g/L at 36, 48 and 72 hours respectively. In the parallel fermentation at 40°C, the cell mass approached 120 optical density units at  $OD_{600}$  and 56 hours, and the pantothenate titers were 18 g/L, 26 g/L and 37 g/L at 36, 48 and 72 hours, respectively.

In another study, PA824 was grown in a fermentor at 43°C in the medium described in Table 23. The cell mass exceeded 160 optical density units at  $OD_{600}$  and 36 hours, and the pantothenate titers were 23 g/L, 34 g/L, 37 g/L and 40 g/L at 24, 36, 48 and 60 hours, respectively. In other fermentations, increasing the amount of trace elements in the glucose feed (e.g., increasing the concentration of SM from 1X to 2X) resulted in even higher titers of pantothenate.

#### **EXAMPLE XII: Identification and characterization of the *B. subtilis coaA* gene product**

The annotated version of the *B. subtilis* genome sequence available on the "Subtilist" web site contains no gene labeled as *coaA*. However a homology search using the protein sequence of *E. coli* pantothenate kinase as a query sequence gave a good match with *B. subtilis* gene *yqiS*, which is annotated as "unknown; similar to pantothenate kinase." This gene appears to be the penultimate gene in an operon containing five open reading frames (Figure 18). Two of the open reading frames encode proteins which are similar to D-serine dehydratase and to "ketoacyl reductase"; the other two have no known homologies. For the open reading frame corresponding to *coaA*, there are three possible start codons; each having a possible ribosome-binding site (RBS) associated with it. The three potential *coaA* ORFs were named *coaA1*, *coaA2*, and *coaA3*, from longest to shortest.



All three potential *coaA* open reading frames were cloned along with their respective RBSs by PCR followed by ligation into expression plasmid pAN229. pAN229 is a low copy vector in *E. coli* that provides expression from the SP01 phage *P<sub>15</sub>* promoter and can integrate by single crossover at *bpr* with tetracycline selection. A representative resulting plasmid, pAN281, is shown in Figure 19.

To determine if the cloned putative *coaA* ORFs actually encode a pantothenate kinase activity, several isolates of all three plasmids were transformed into the *E. coli* strain YH1, that contains the *coaA15(Ts)* allele. Transformants were streaked to plates incubated at 30° and 43°C to test for complementation of the temperature sensitive allele. All isolates of all three *coaA* variants, except for one isolate of pAN282, complemented well at 43°C, indicating that all three plasmid constructs encode an active pantothenate kinase. Accordingly, it can be concluded that the *B. subtilis yqjS* open reading frame codes for an active pantothenate kinase.

#### 15 **EXAMPLE XIII: Deletion of the *coaA* gene from the *B. subtilis* genome**

The *coaA* gene of *B. subtilis* (*yqjS*) was deleted from the chromosome of a *B. subtilis* strain by conventional means. The majority of the *coaA* coding sequence was deleted from a plasmid clone and replaced by a chloramphenicol resistance gene (*cat*), while leaving approximately 1 kb of upstream and downstream sequence to allow homologous recombination within the chromosome, to give plasmid pAN296 (see Figure 17). pAN296 was then used to transform a *B. subtilis* strain (PY79), selecting for chloramphenicol resistance. The majority of transformants result from a double crossover event that effectively substitutes the *cat* gene for the *coaA* gene. The transformed strain containing the *coaA* deletion – *cat* insertion grew normally due the presence of a second *B. subtilis* pantothenate kinase encoding gene described herein.

#### **EXAMPLE XIV: Identification and characterization of a second *B. subtilis* gene encoding pantothenate kinase activity**

As described in detail in the instant specification, in order to maximize pantothenate production, it is necessary to restrict the flow of pantothenate toward Coenzyme A (CoA), for example, by reducing the activity of pantothenate kinase, the first enzyme in the pathway from pantothenate to CoA. After finding that deletion of the *coaA* gene from the chromosome of *B. subtilis* is not a lethal event (see Example XIII), it was concluded that *B. subtilis* must contain a second gene that encodes an active pantothenate kinase, since pantothenate kinase is an essential enzyme activity.

A second pantothenate kinase-encoding gene was identified by complementing the *E. coli* strain YH1 (*coaA15(Ts)*) with a *B. subtilis* gene bank and selecting for transformants that were able to grow at 43°C. Found among the transformants were two families of plasmids that had overlapping restriction maps within each family, but not  
5 between the families. As expected, the restriction map of one family was identical to that predicted from the *B. subtilis* genome sequence for the homologue of the *E. coli* *coaA* gene (which we named *coaA* also, see above) and surrounding sequences. The other family had a restriction map that was completely non-overlapping with the first.

DNA sequencing of the ends of the cloned inserts from the second family  
10 showed that the clones came from a region of the *B. subtilis* chromosome that includes the 3' end of the *ftsH* gene, the 5' end of the *sul* gene, and all of the *yacB*, *yacC*, *yacD*, *cysK*, *pabB*, *pabA* and *pabC* genes. None of the open reading frames of these cloned inserts showed homology to any known pantothenate kinase sequences, either prokaryotic or eukaryotic.

Several deletions were created through the *B. subtilis* genomic sequences in the  
15 cloned inserts. Each deletion was tested for complementation of the *E. coli* temperature sensitive pantothenate kinase. In particular, a deletion that removed all DNA between a *Stu* I site in the cloning vector and a *Swa* I site in the *yacC* gene, leaves *yacB* as the only intact open reading frame in the cloned insert (see Figure 21). This deleted plasmid still  
20 complemented the *E. coli* pantothenate kinase mutant. However, another deletion that removed DNA from the *Swa* I site in *yacC* through a *Bst*1107I site in the (already truncated) *ftsH* gene, could not complement the *E. coli* pantothenate kinase mutant. From these results, it was concluded that the *yacB* open reading frame was responsible for the complementation activity. To confirm that *yacB* is a pantothenate kinase gene,  
25 the *yacB* ORF plus 112 base pairs of downstream flanking sequence was amplified by PCR in two independent reactions and cloned downstream of a constitutive promote to give plasmids pAN341 and pAN342 (Figure 22). Both pAN341 and pAN342 complemented the defect in YH1 at 44°C, while a control plasmid, which has the same backbone, but expresses *panBCD* instead of *yacB* did not. This confirmed that the *yacB*  
30 open reading frame was responsible for the complementation of YH1.

As such, a novel gene that encodes pantothenate kinase activity in *B. subtilis* has been discovered that is not related by homology to any previously known pantothenate kinase gene. This gene has been renamed *coaX*, as a second, alternative gene that encodes an enzyme that catalyzes the first step in the pathway from pantothenate to  
35 CoaA. Deletion of *coaX* by methods described above for deleting *coaA*, in conjunction

with reduction in the activity of the CoaA enzyme, provides a means to reduce pantothenate kinase activity to the desired level.

Several homologues of the *B. subtilis coaX* gene were identified by homology searching of various publically available databases using the published *yacB* (*coaX*) open reading frame sequence and predicted amino acid sequence (as set forth in SEQ ID NOs:84 and 85 respectively). In two cases (*Mycobacterium tuberculosis* and *Streptomyces coelicolor*) the homologous *coaX* genes are adjacent to, or almost adjacent to, pantothenate biosynthetic genes, consistent with these homologs having a role in pantothenate metabolism. The CoaX proteins show no homology to the CoaA family of pantothenate kinases, nor to the eukaryotic family of pantothenate kinases exemplified by PanK of *Saccharomyces cerevisiae*.

Alignment of the amino acid sequences of several bacterial CoaX homologs with the amino acid sequence predicted from translating the *B. subtilis yacB* ORF described in the published *B. subtilis* genome sequence revealed that the CoaX proteins from other bacteria contained additional amino acid residues at their carboxy-terminal ends. Moreover, these extensions beyond the end of the predicted amino acid for the *B. subtilis* gene product contained two relatively well conserved segments of sequence.

Translation of nucleotide sequences just downstream from the stop codon of the *B. subtilis yacB* ORF in a different reading frame revealed the existence of amino acid sequences very similar to the carboxy-terminal extensions of the other bacterial CoaX proteins. It is thus believed that an error exists in the published DNA sequence of the *B. subtilis yacB* ORF sequence that causes a frame shift leading to an artifactual downstream amino acid sequence and premature termination.

The PCR-generated sequences of *B. subtilis CoaX* in pAN341 and pAN342 (described above) contain enough downstream flanking sequence to encode the putative carboxy-terminal extension described above, which is consistent with the result that the clones were functional in the complementation assay. However when the 3' PCR primer was positioned to include only the shorter *yacB* ORF predicted from the published sequence, but not to include the putative carboxy-terminal extension, then the resulting plasmids, pAN329 and pAN330 (similar in structure to pAN341 and pAN342; see Figure 22), did not complement the defect in YH1. This result supports the notion that the published *yacB* coding sequence contains a frame-shift error, and that the carboxy-terminal end of CoaX is necessary for pantothenate kinase activity. The predicted correct nucleotide sequence for *B. subtilis coaX* is set forth as SEQ ID NO:19 and the translated amino acid sequence is set forth as SEQ ID NO:9. A multiple

**EXAMPLE XVIII: Generation of mutant *coaX* genes encoding pantothenate kinase having reduced or temperature sensitive activities**

Mutant *coaX* genes are generated by introducing point mutations into the gene and testing the resulting mutants for the ability to complement the *E. coli* YH1 strain as described in Example XII. Preferred mutations in the *coaX* gene sequences are those that encode a substitution of a residue conserved among CoaX sequences from a variety of bacterial sources (e.g., a conserved residue set forth in Figure 23). Alternatively, random mutations in the *coaX* gene sequence are generated by mutagenic PCR and *in vitro* recombination and identified by screening for alleles that poorly complement the *E. coli coaA15(Ts)* mutant.

Mutants so generated (i.e., mutants having reduced *coaX* activity) can be further engineered such that the endogenous *coaA* gene is deleted (as described in Example XIII). CoaX reduced-activity mutants can also be further engineered to contain reduced-activity CoaA gene products as described in Example XV.

15

**EXAMPLE XIX: Enhanced Production of Panto-Compounds Using Bacteria Having Deletions in One or More Pantothenate Biosynthetic Enzymes**

If the desired panto-compound is not pantothenate, then an appropriate deletion of one or more of the pantothenate biosynthetic genes from a pantothenate overproducing strain will provide a strain that produces said desired panto-compound. In this example, the desired panto-compound is pantoate. Starting with, for example, strain PA236, PA313 or PA824 either one or both of the *panC* and *panD* genes is deleted. In another example, ketopantoate is the desired panto-compound. Starting with, for example strain PA244, PA245 or PA824 one, two or all of the *ilvC*, *panE1*, *panC* and *panD* genes are deleted from the starting strain. If  $\beta$ -alanine is the desired panto-compound, then *panB* and *panC* can be deleted, preferably in a fashion that leaves an in frame fusion of a small portion of the 5' end of *panB* with a small portion of the 3' end of *panC*, from the strain PA221, PA235, PA245, or PA313. In all of the above-mentioned examples, the panto-compound producing strain will be a pantothenate auxotroph. Accordingly, the growth medium requires sufficient pantothenate for adequate growth. Vectors designed to overexpress *panD* as described above are then transformed into the above strains to further enhance  $\beta$ -alanine production.

The above-mentioned deletions are accomplished by methods well-known to those skilled in the art, for example, by insertion of an antibiotic resistance gene and removing sufficient sequence from the target gene(s) to inactivate said target gene(s).

35

Alternatively, removal of targeted sequences is accomplished without simultaneous introduction of an antibiotic resistance gene in said target gene and then introduced by congression (co-transformation with any other appropriate selectable DNA sequence) followed by screening for the loss of function of said target gene by replica plating.

5

Table 24 : Strains (and corresponding phenotypes) for panto-compound production

Name	Pheno type	Drug resist.	<i>panBCD</i> locus	<i>panE</i> locus	<i>ilvD</i> locus	<i>amyE</i> locus	<i>bpr</i> locus	Parent
PA221	Trp-		<i>P26panBCD</i>					
PA222			<i>P<sub>15</sub> panBCD</i>					RL-1
PA235			<i>P26panBCD</i>					
PA236			<i>P<sub>26</sub> panBCD</i>	<i>P<sub>26</sub> panE1</i>				PA221
PA327	Trp-		<i>P26panBCD</i>	<i>P26panE1</i>				PA221
PA328	Trp-		<i>P26panBCD</i>	<i>P26panE1</i>				PA235
PA340	Trp-	Spc	<i>P26panBCD</i>	<i>P26panE1</i>		<i>P26ilvBNC</i>		PA327
PA342	Trp-	Spc	<i>P26panBCD</i>	<i>P26panE1</i>		<i>P26ilvBNC</i>		PA328
PA354	Trp-	Spc	<i>P26panBCD</i>	<i>P26panE1</i>	<i>P26ilvD</i>	<i>P26ilvBNC</i>		PA342
PA365	Trp-	Spc, Tet	<i>P26panBCD</i>	<i>P26panE1</i>	<i>P26ilvD</i>	<i>P26ilvBNC</i>	<i>P26panD423</i>	PA354
PA374	Trp-	Spc	<i>P26panBCD</i>	<i>P26panE1</i>	<i>P26ilvD</i>	<i>P26ilvBNC</i>		PA340
PA377	Trp-	Spc, Tet	<i>P26panBCD</i>	<i>P26panE1</i>	<i>P26ilvD</i>	<i>P26ilvBNC</i>	<i>P26panD423</i>	PA374
PA401	Trp-		<i>P26panBCD</i>				<i>P26panD423</i>	PA221
PA402	Trp-		<i>P26panBCD</i>				<i>P26panD428</i>	PA221
PA403	Trp-		<i>P26panBCD</i>				<i>P26panD429</i>	PA221
PA404	Trp-	Spc, Tet	<i>P26panBCD</i>	<i>P26panE1</i>		<i>P26ilvBNC</i>	<i>P26panD423</i>	PA340
PA405	Trp-	Spc, Tet	<i>P26panBCD</i>	<i>P26panE1</i>		<i>P26ilvBNC</i>	<i>P26panD423</i>	PA342
PA651	Trp-	Spc	<i>P26panBC*D</i>	<i>P26panE1</i>	<i>P26ilvD</i>	<i>P26ilvBNC</i>		PA374
PA284		Spc, Tet	<i>P26'panBCD</i>	<i>P26panE1</i>	<i>P26ilvD</i>	<i>P26ilvBNC</i>	<i>P26panD423</i>	PA377

**Equivalents** Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

10

**What is claimed:**

1. A method of producing a panto-compound comprising culturing a microorganism which overexpresses at least one *Bacillus* pantothenate biosynthetic enzyme under conditions such that the panto-compound is produced.
2. The method of claim 1, wherein the microorganism overexpresses at least one *Bacillus subtilis* pantothenate biosynthetic enzyme.
3. The method of claim 1 or 2, wherein the pantothenate biosynthetic enzyme is selected from the group consisting of ketopantoate hydroxymethyltransferase, pantothenate synthetase, aspartate- $\alpha$ -decarboxylase and ketopantoate reductase.
4. The method of any one of claims 1 to 3, wherein the microorganism overexpresses at least two pantothenate biosynthetic enzymes.
5. The method of any one of claims 1 to 3, wherein the microorganism overexpresses at least three pantothenate biosynthetic enzymes.
6. The method of any one of claims 1 to 5, wherein the panto-compound is selected from the group consisting of pantothenate, pantoate, ketopantoate and  $\beta$ -alanine.
7. A method of producing a panto-compound comprising culturing a ketopantoate reductase-overexpressing (KPAR-O) microorganism under conditions such that the panto-compound is produced.
8. The method of claim 7, wherein the panto-compound is pantothenate or pantoate.
9. The method of claim 7 or 8, wherein the ketopantoate reductase is bacterial-derived.
10. The method of claim 7 or 8, wherein the ketopantoate reductase is derived from *Bacillus*.

11. The method of claim 7 or 8, wherein the ketopantoate reductase is derived from *Bacillus subtilis*.

5 12. The method of any one of claims 7 to 11, wherein the KPAR-O microorganism further overexpresses at least one pantothenate biosynthetic enzyme in addition to overexpressing ketopantoate reductase.

10 13. The method of claim 12, wherein the KPAR-O microorganism further overexpresses at least one of ketopantoate hydroxymethyltransferase, pantothenate synthetase and aspartate- $\alpha$ -decarboxylase.

14. A method of producing pantothenate in a manner independent of precursor feed comprising culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism having a deregulated isoleucine-valine (*ilv*) pathway under conditions  
15 such that pantothenate is produced.

15. A method of producing at least 2 g/L pantothenate in a manner independent of aspartate or  $\beta$ -alanine feed comprising culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism under conditions such that  
20 pantothenate is produced.

16. A method of producing at least 2 g/L pantothenate in a manner independent of valine or  $\alpha$ -ketoisovalerate feed comprising culturing a microorganism having a deregulated isoleucine-valine (*ilv*) biosynthetic pathway under conditions such  
25 that pantothenate is produced.

17. A method of producing at least 30 g/L pantothenate in a manner independent of aspartate or  $\beta$ -alanine feed comprising culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism under conditions such that  
30 pantothenate is produced.

18. A method of producing at least 30 g/L pantothenate in a manner independent of valine or  $\alpha$ -ketoisovalerate feed comprising culturing a microorganism having a deregulated isoleucine-valine (*ilv*) biosynthetic pathway under conditions such  
35 that pantothenate is produced.

19. A  $\beta$ -alanine independent high yield production method for producing pantothenate comprising culturing a manipulated microorganism under conditions such that pantothenate is produced at a significantly high yield.

5 20. The method of any one of claims 14 to 19, wherein the microorganism overexpresses acetohydroxyacid synthetase or is transformed with a vector comprising an *ilvBN* nucleic acid sequence or an *alsS* sequence.

10 21. The method of any one of claims 14 to 19, wherein the microorganism overexpresses acetohydroxyacid isomeroeductase or is transformed with a vector comprising an *ilvC* nucleic acid sequence.

15 22. The method of any one of claims 14 to 19, wherein the microorganism overexpresses dihydroxyacid dehydratase or is transformed with a vector comprising an *ilvD* nucleic acid sequence.

20 23. The method of any one of claims 19 to 22, wherein the microorganism overexpresses aspartate- $\alpha$ -decarboxylase or is transformed with a vector comprising a *panD* nucleic acid sequence.

24. The method of any one of claims 14 to 23, wherein the microorganism further has a deregulated pantothenate biosynthetic pathway.

25 25. The method of any one of claims 14 to 24, wherein the microorganism further has at least one mutant gene selected from the group consisting of a mutant *avtA* gene, a mutant *ilvE* gene, a mutant *ansB* gene and a mutant *alsD* gene.

30 26. The method of claim 24, wherein the microorganism overexpresses any of ketopantoate hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase and aspartate- $\alpha$ -decarboxylase.

35 27. The method of claim 24 or 26, wherein the microorganism is transformed with a vector comprising a *panBCD* nucleic acid sequence or a vector comprising a *panE1* nucleic acid sequence.



28. The method of any one of claims 14 to 16, and 19 to 27, wherein pantothenate is produced at a level selected from the group consisting of a level greater than 10g/L, a level greater than 20g/L and a level greater than 40g/L.

29. The method of claim 20, wherein the microorganism overexpresses acetohydroxyacid synthetase derived from *Bacillus* or is transformed with a vector comprising an *ilvBN* nucleic acid sequence or an *alsS* nucleic acid sequence derived from *Bacillus*.

30. The method of claim 21, wherein the microorganism overexpresses acetohydroxyacid isomeroreductase derived from *Bacillus* or is transformed with a vector comprising an *ilvC* nucleic acid sequence derived from *Bacillus*.

31. The method of claim 22, wherein the microorganism overexpresses dihydroxyacid dehydratase derived from *Bacillus* or is transformed with a vector comprising an *ilvD* nucleic acid sequence derived from *Bacillus*.

32. The method of claim 23, wherein the microorganism overexpresses aspartate- $\alpha$ -decarboxylase derived from *Bacillus* or is transformed with a vector comprising a *panD* nucleic acid sequence derived from *Bacillus*.

33. The method of claim 24 or 26, wherein the microorganism overexpresses any of ketopantoate hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase and aspartate- $\alpha$ -decarboxylase derived from *Bacillus*.

34. The method of claim 27, wherein the vector comprises a *panBCD* nucleic acid sequence or a *panE1* nucleic acid sequence derived from *Bacillus*.

35. A method of producing a panto-compound comprising contacting a composition comprising at least one pantothenate biosynthesis pathway precursor or isoleucine-valine biosynthesis pathway precursor with at least one isolated *Bacillus* enzyme selected from the group consisting of ketopantoate hydroxymethyltransferase, ketopantoate reductase, pantothenate synthetase and aspartate- $\alpha$ -decarboxylase, under conditions such that the panto-compound is produced.

36. A method of producing  $\beta$ -alanine comprising culturing an aspartate- $\alpha$ -decarboxylase-overexpressing (A $\alpha$ D-O) microorganism under conditions such that  $\beta$ -alanine is produced.

5        37. The method of claim 36, wherein the A $\alpha$ D-O microorganism has a mutation in a nucleic acid sequence encoding a pantothenate biosynthetic enzyme selected from the group consisting of ketopantoate hydroxymethyltransferase, ketopantoate reductase and pantothenate synthetase.

10       38. A method of producing  $\beta$ -alanine comprising contacting a composition comprising aspartate with an isolated *Bacillus* aspartate- $\alpha$ -decarboxylase enzyme under conditions such that  $\beta$ -alanine is produced.

15       39. A method for enhancing production of a panto-compound comprising culturing a mutant microorganism having a mutant *coaX* gene under conditions such that the panto-compound production is enhanced.

20       40. The method of claim 39, wherein said recombinant microorganism has a mutant *coaA* gene.

      41. A method of producing a panto-compound comprising a pantothenate kinase mutant microorganism under conditions such that the panto-compound is produced at a significantly high yield.

25       42. The method of claim 41, wherein said mutant microorganism has a mutant *coaA* gene.

      43. The method of claim 41, wherein said mutant microorganism has a mutant *coaX* gene.

30       44. The method of claim 41, where said mutant microorganism has a mutant *coaA* and *coaX* gene.

35       45. The method of any one of claims 39 to 44, wherein said panto-compound is selected from the group consisting of ketopantoate, pantoate or pantothenate.

46. The method of any one of claims 39 to 44, wherein said panto-compound is pantothenate.

47. The method of any one of claims 39 to 44, wherein said panto-compound is produced at a level selected from the group consisting of a level greater than 10g/L, a level greater than 20g/L and a level greater than 40g/L.

48. The method of any one of claims 39 to 44, wherein said recombinant microorganism further has a deregulated pantothenate biosynthetic pathway or further has a deregulated isoleucine-valine (*ilv*) biosynthetic pathway.

49. The method of claim any one of claims 39 to 44, wherein said recombinant microorganism further overexpresses *panD* and *panE*.

50. The method of any one of claims 39 to 44, wherein said recombinant microorganism further has at least one mutant gene selected from the group consisting of a mutant *avtA* gene, a mutant *ilvE* gene, a mutant *ansB* gene and a mutant *alsD* gene.

51. A method for enhancing production of a panto-compound comprising culturing a microorganism that has a deregulated pantothenate biosynthetic pathway and that also has a mutation that results in reduced pantothenate kinase activity under conditions such that the panto-compound production is enhanced.

52. A method for identifying compounds which modulate pantothenate kinase activity comprising contacting a recombinant cell expressing pantothenate kinase encoded by the *coaX* gene with a test compound and determining the ability of the test compound to modulate pantothenate kinase activity in said cell.

53. The method of claim 52, wherein said cell further comprises a mutant *coaA* gene encoding a pantothenate kinase having reduced activity.

54. The method of any one of claims 1 to 51, wherein the microorganism is Gram positive.

55. The method of any one of claims 1 to 51, wherein the microorganism is Gram negative.

56. The method of any one of claims 1 to 51, wherein the microorganism is a microorganism belonging to a genus selected from the group consisting of *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* and *Streptomyces*.

5 57. The method of any one of claims 1 to 51 and 54 to 56, wherein the microorganism is of the genus *Bacillus*.

58. The method of any one of claims 1 to 51 and 54 to 57, wherein the microorganism is *Bacillus subtilis*.

10

59. The method of any one of claims 1 to 13, 35, 39 to 51 and 54 to 58, further comprising recovering the panto-compound.

15

60. The method of any one of claims 14 to 34 and 54 to 58, further comprising recovering the pantothenate.

61. The method of any one of claims 1 to 14, 35, 39 to 46, 48 to 51 and 54 to 59, wherein the panto-compound is produced at a level greater than 2 g/L.

20

62. A recombinant microorganism which overexpresses at least one *Bacillus* pantothenate biosynthetic enzyme.

25

63. The recombinant microorganism of claim 62, which overexpresses at least one *Bacillus subtilis* pantothenate biosynthetic enzyme.

64. The recombinant microorganism of claim 62 or 63, wherein the pantothenate biosynthetic enzyme is selected from the group consisting of ketopantoate hydroxymethyltransferase, pantothenate synthetase, aspartate- $\alpha$ -decarboxylase and ketopantoate reductase.

30

65. The recombinant microorganism of any one of claims 62 to 64, wherein the pantothenate biosynthetic enzyme is ketopantoate reductase.

66. A recombinant microorganism which overexpresses aspartate- $\alpha$ -decarboxylase and has a deregulated isoleucine-valine (*ilv*) biosynthetic pathway.

35

67. A recombinant microorganism having a mutant *coaX* gene, said mutant *coaX* gene encoding reduced pantothenate kinase activity in said microorganism.

5 68. The recombinant microorganism of claim 67 further having a mutant *coaA* gene, said mutant *coaA* gene encoding reduced pantothenate kinase activity in said microorganism.

69. A recombinant microorganism having a mutant *coaX* gene and optionally having a mutant *coaA* gene, said mutant microorganism having reduced pantothenate  
10 kinase activity as compared to a microorganism having wild-type *coaA* and *coaX* genes.

70. A recombinant microorganism comprising a vector comprising an isolated *coaX* gene.

15 71. A recombinant microorganism that overproduces a panto-compound, the microorganism having a deregulated pantothenate biosynthetic pathway and having at least one mutation that results in a decrease in the capacity of the microorganism to synthesize Coenzyme A (CoA).

20 72. The recombinant microorganism of claim 71, having at least one mutation that results in a reduced level of pantothenate kinase activity.

73. The recombinant microorganism of claim 72, having a mutation in a *coaA* gene, or homologue thereof, that results in a reduced level of CoaA enzyme  
25 activity.

74. The recombinant microorganism of claim 72, having a mutation in a *coaX* gene, or homologue thereof, that results in a reduced level of CoaX enzyme  
30 activity.

75. The recombinant microorganism of claim 72, having a mutation in a *coaA* gene, or homologue thereof, and having a mutation in a *coaX* gene, or homologue thereof, the mutations resulting in reduced levels of CoaA enzyme activity and reduced  
35 CoaX enzyme activity.

76. The recombinant microorganism of any one of claims 66 to 70 which further has a deregulated pantothenate biosynthetic pathway.

77. The recombinant microorganism of any one of claims 62 to 65 and 67 to 75, further having a deregulated isoleucine-valine (*ilv*) biosynthetic pathway.
- 5 78. The recombinant microorganism of any one of claims 62 to 77, which is Gram positive.
79. The recombinant microorganism of claim 78 belonging to a genus selected from the group consisting of *Bacillus*, *Corynebacterium*, *Lactobacillus*,  
10 *Lactococci* and *Streptomyces*.
80. The recombinant microorganism of claim 79 belonging to the genus *Bacillus*.
- 15 81. The recombinant microorganism of claim 80 which is *Bacillus subtilis*.
82. A recombinant microorganism selected from the group consisting of PA221, PA235, PA236, PA313, PA410, PA402, PA403, PA411, PA412, PA413, PA303, PA327, PA328, PA401, PA340, PA342, PA404, PA405, PA374, PA354,  
20 PA365, PA377, PA651 and PA824.
83. A recombinant vector for use in the production of panto-compounds comprising a nucleic acid sequence which encodes at least one *Bacillus* pantothenate biosynthetic enzyme operably linked to regulatory sequences.
- 25 84. The vector of claim 83, comprising a nucleic acid sequence which encodes at least one *Bacillus subtilis* pantothenate biosynthetic enzyme.
85. The vector of claim 84, wherein the nucleic acid sequence encodes at  
30 least one of ketopantoate hydroxymethyltransferase, pantothenate synthetase, aspartate- $\alpha$ -decarboxylase and ketopantoate reductase.
86. A recombinant vector comprising at least one nucleic acid sequence selected from the group consisting of SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27,  
35 SEQ ID NO:29 and SEQ ID NO:59.

87. The vector of claim 84, wherein the nucleic acid sequence encodes ketopantoate reductase.

88. A vector comprising a mutant *coaX* gene, said mutant encoding a  
5 pantothenate kinase enzyme having reduced activity.

89. A vector comprising an isolated *coaX* gene.

90. A vector comprising an isolated *Bacillus coaX* gene.

91. A vector comprising an isolated *Bacillus subtilis coaX* gene.

92. The vector of any one of claims 86 and 89 to 91, which further comprises regulatory sequences.

93. The vector of any one of claims 83 to 85, 87 and 92, wherein the regulatory sequences comprise a constitutively active promoter.

94. The vector of claim 93, wherein the constitutively active promoter  
20 comprises  $P_{veg}$  (SEQ ID NO:41),  $P_{15}$  (SEQ ID NO:39) or  $P_{26}$  (SEQ ID NO:40) sequences.

95. The vector of claim 83, wherein the regulatory sequences comprise at least one artificial ribosome binding site (RBS).

96. The vector of claim 95, wherein the artificial RBS comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56 and SEQ ID NO:57.

97. A vector selected from the group consisting of pAN004, pAN005, pAN006, pAN236, pAN423, pAN428, pAN429, pAN441, pAN442, pAN443, pAN251, pAN267, pAN256, pAN257, pAN263, pAN240, pAN294, pAN296, pAN336, pAN341 and pAN342.

98. A recombinant microorganism comprising the vector of claim 86 or 93.

99. An isolated nucleic acid molecule which encodes at least one *Bacillus* pantothenate biosynthetic gene.

100. The isolated nucleic acid molecule of claim 99 which encodes at least  
5 one *Bacillus subtilis* pantothenate biosynthetic gene.

101. The isolated nucleic acid molecule of claim 99 or 100 which encodes ketopantoate reductase.

102. An isolated *Bacillus* pantothenate biosynthetic enzyme polypeptide.  
10

103. An isolated *Bacillus subtilis* pantothenate biosynthetic enzyme  
polypeptide.

104. An isolated *Bacillus* ketopantoate reductase polypeptide.  
15

105. An isolated *Bacillus subtilis* ketopantoate reductase polypeptide.

106. An isolated *Bacillus* aspartate- $\alpha$ -decarboxylase polypeptide.  
20

107. An isolated *Bacillus subtilis* aspartate- $\alpha$ -decarboxylase polypeptide.

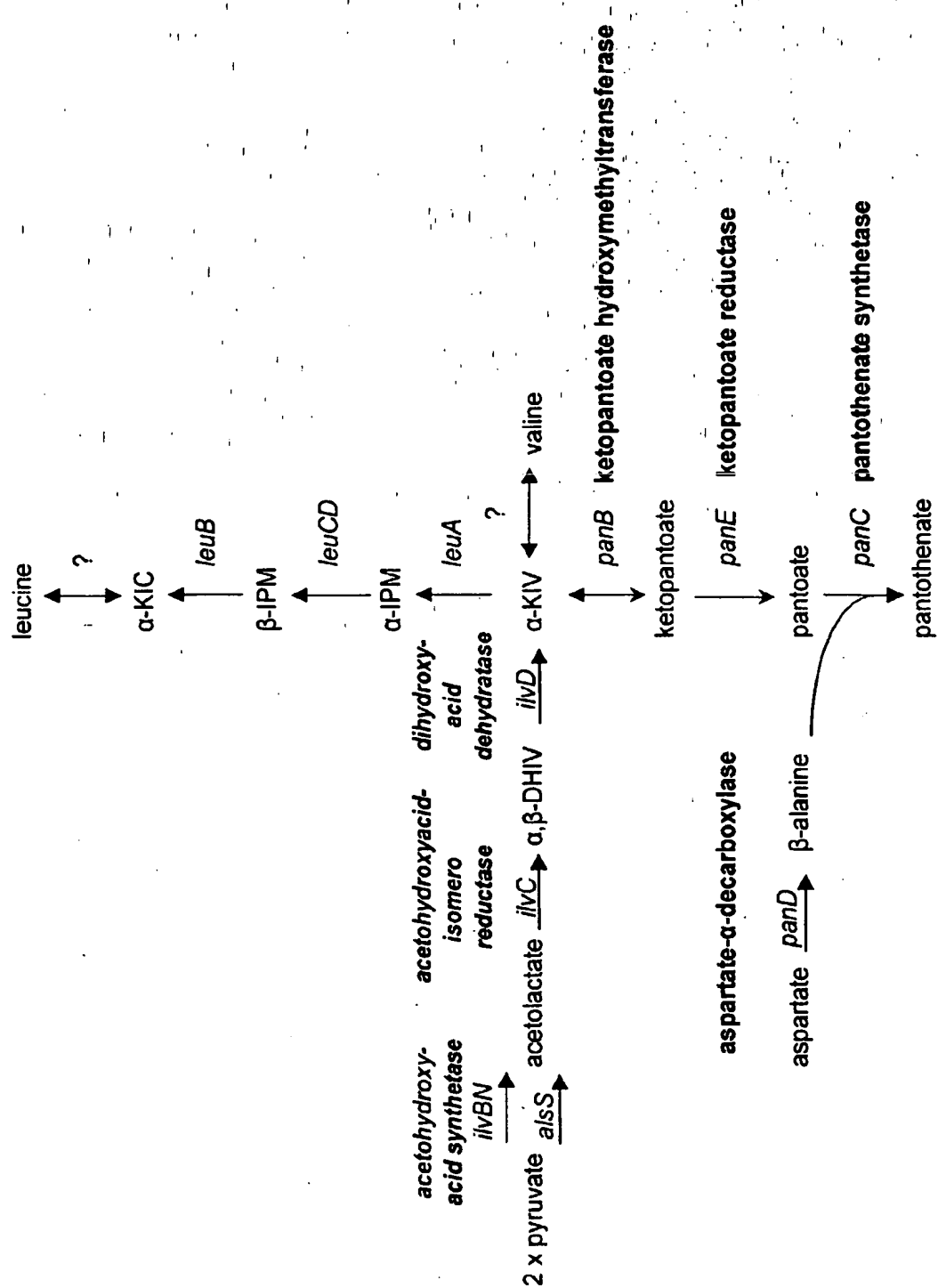
108. An isolated nucleic acid molecule comprising a mutant *coaX* gene.

109. An isolated nucleic acid molecule comprising a *coaX* gene.  
25

110. An isolated pantothenate kinase protein encoded by a *coaX* gene.



FIG. 1



**Figure 2. Plasmid pAN240, containing sequences ligated upstream of the P<sub>26</sub>panBCD cassette, equivalent to the integrated version in strain PA221.**

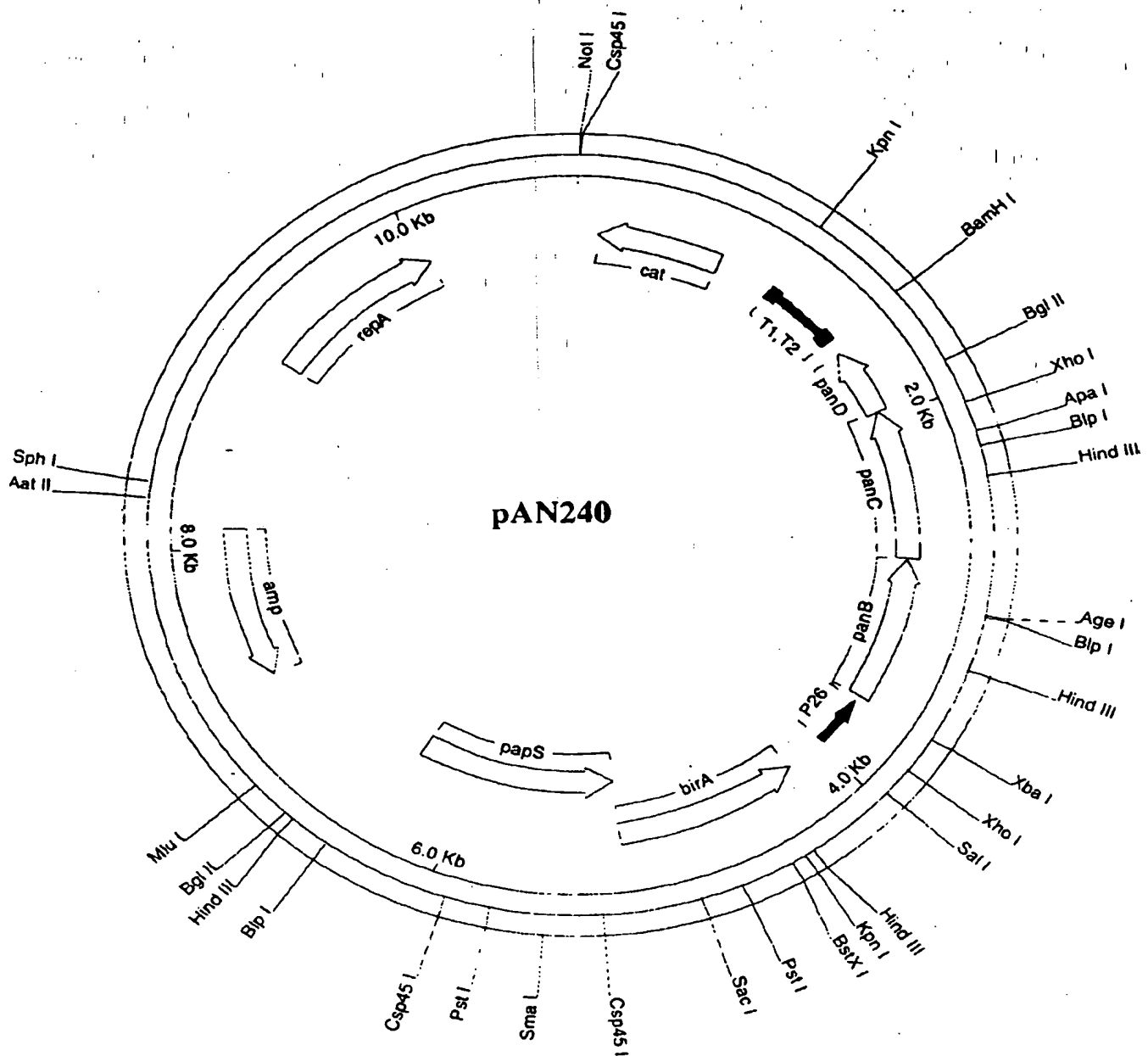


Figure 3A Plasmid pAN004, containing the panBCD operon expressed from P26 and RBS1.

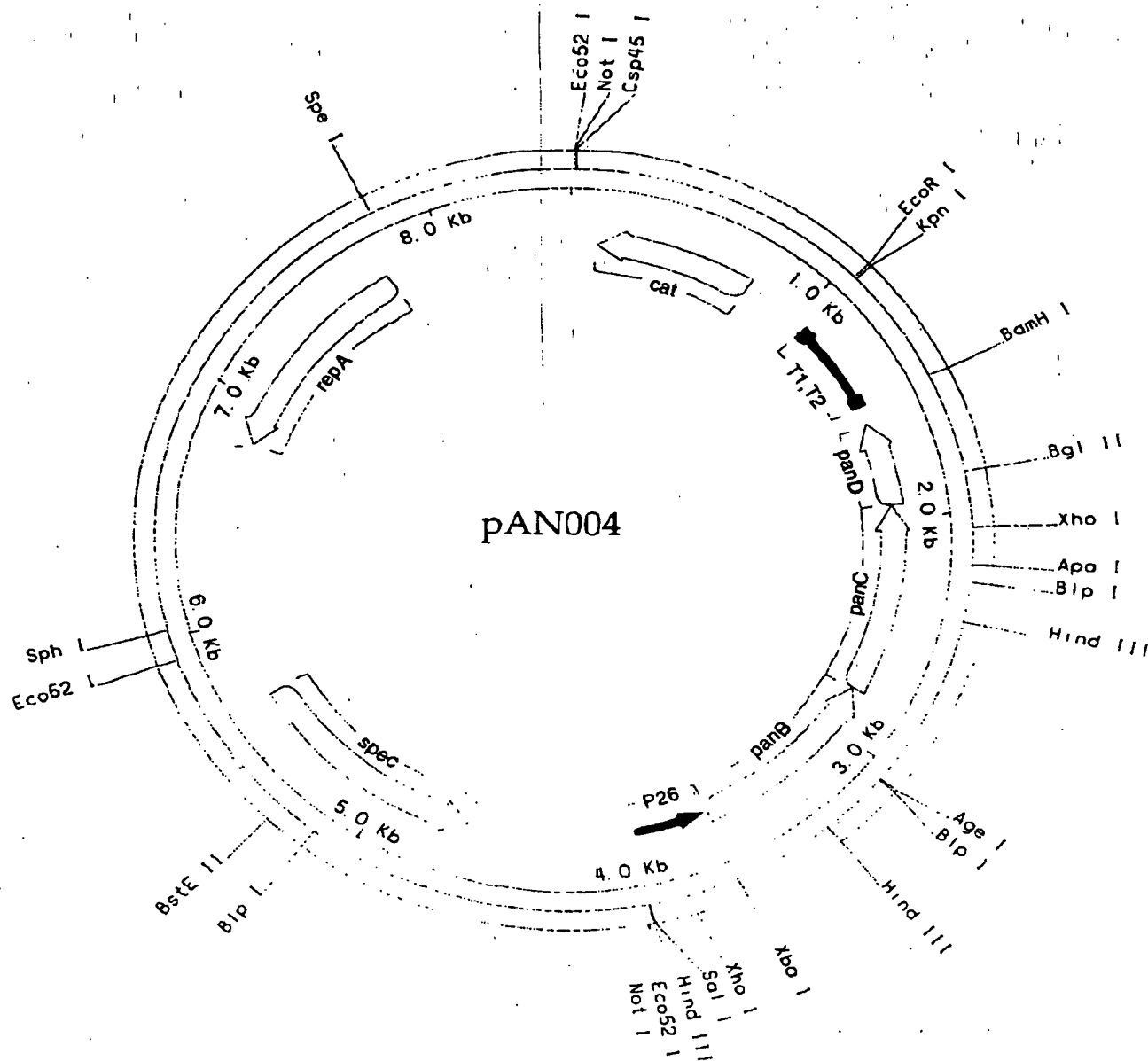
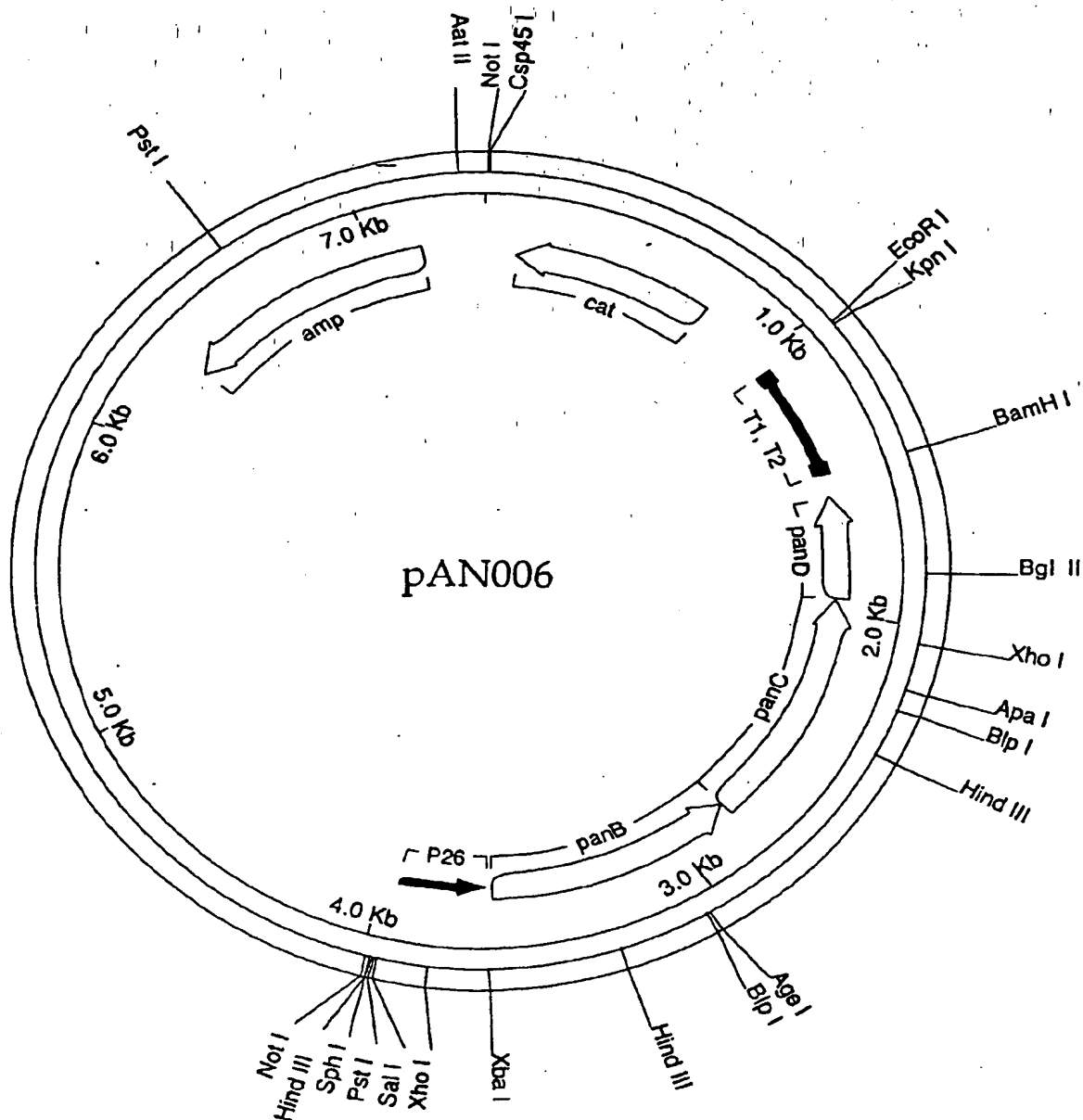


Figure 3 Plasmid pAN006, containing the panBCD operon expressed from P26 and RBS2.



**Figure 4** Plasmid pAN236, containing an integratable and amplifiable P26-RBS2-panE1 expression cassette.

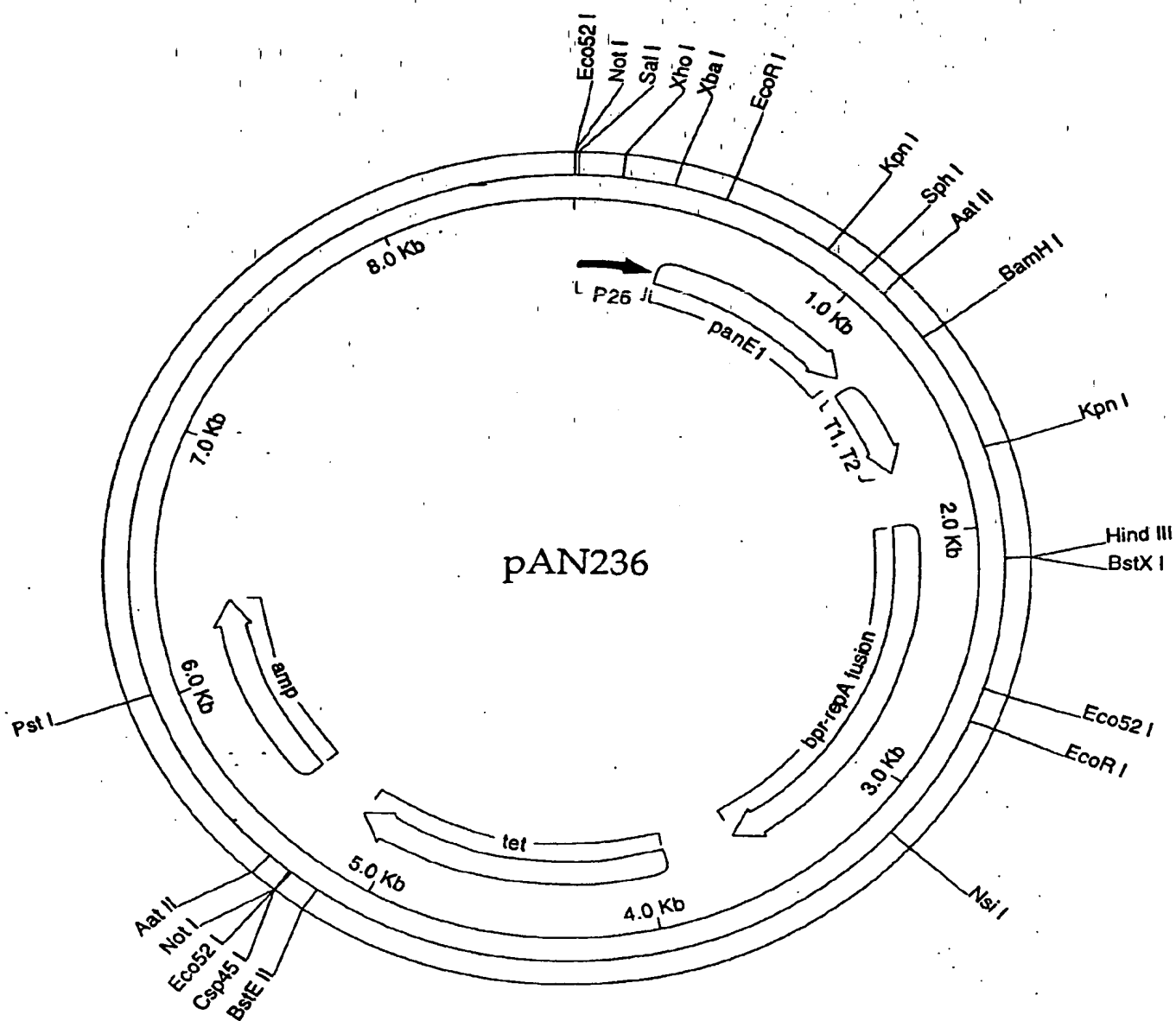


Figure 5 Construction of pAN423

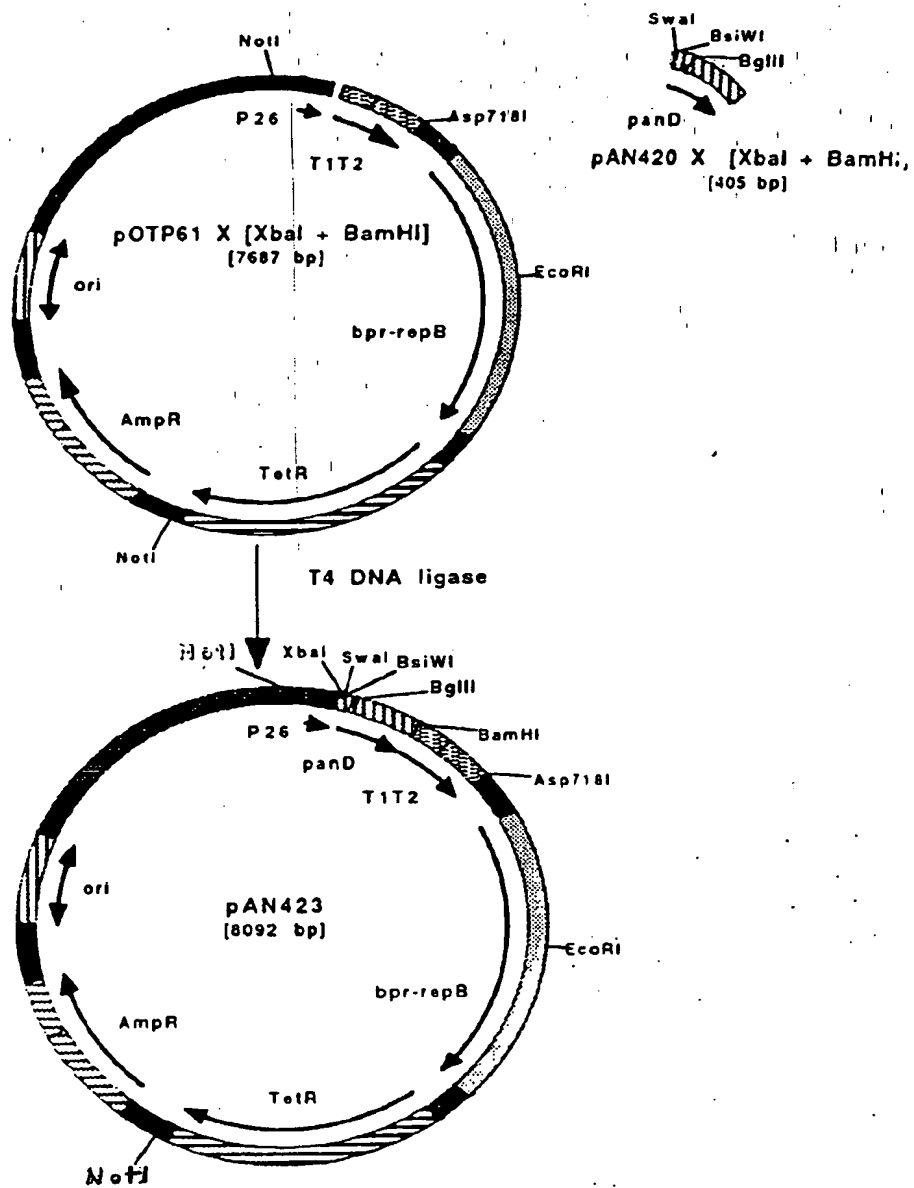


Figure 6 Construction of pAN426 and pAN427.

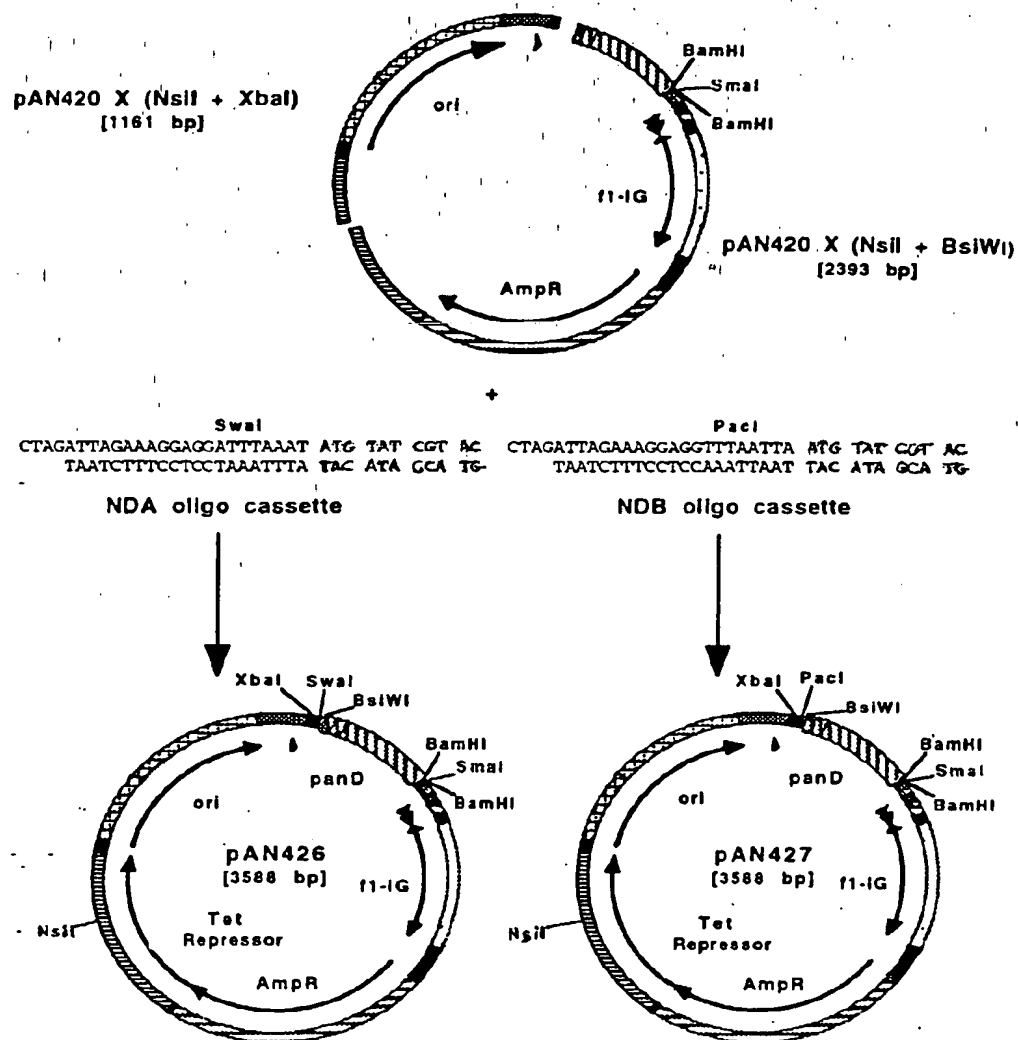
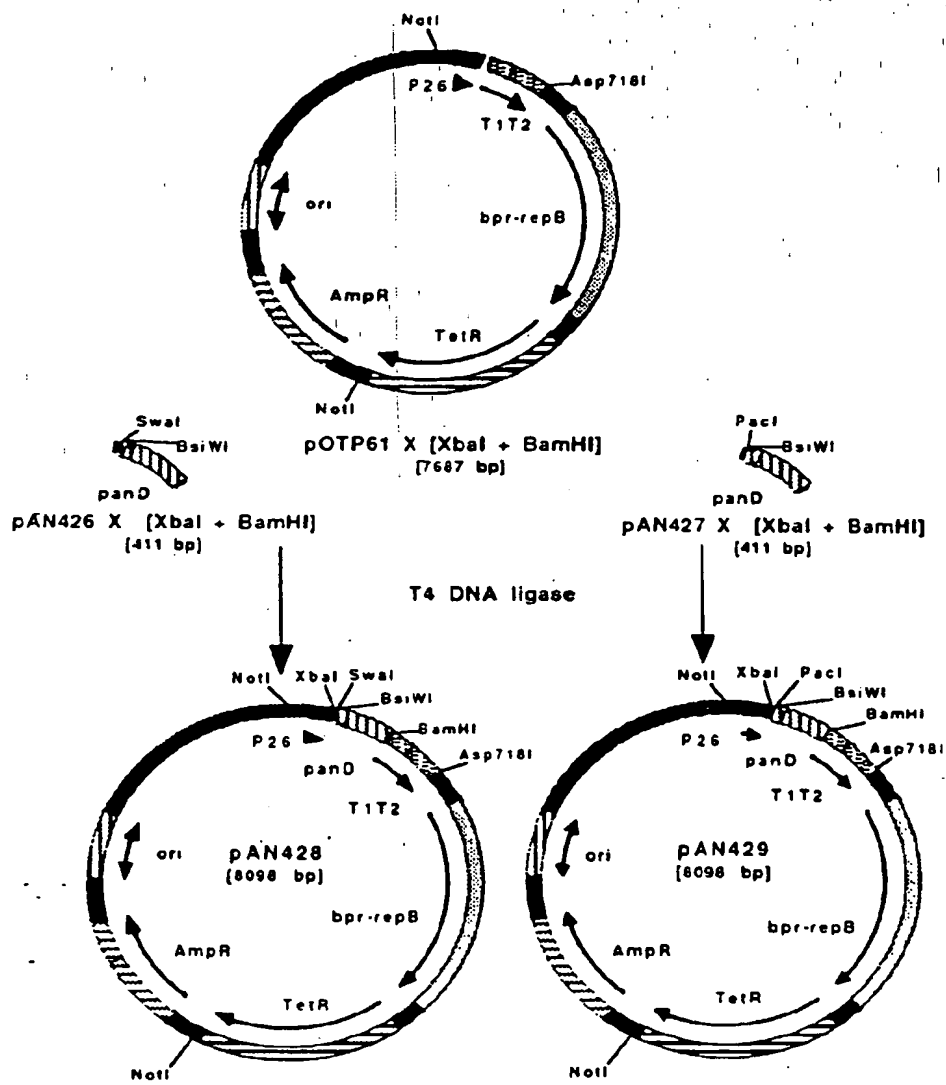


Figure 7 Construction of pAN428 and pAN429.





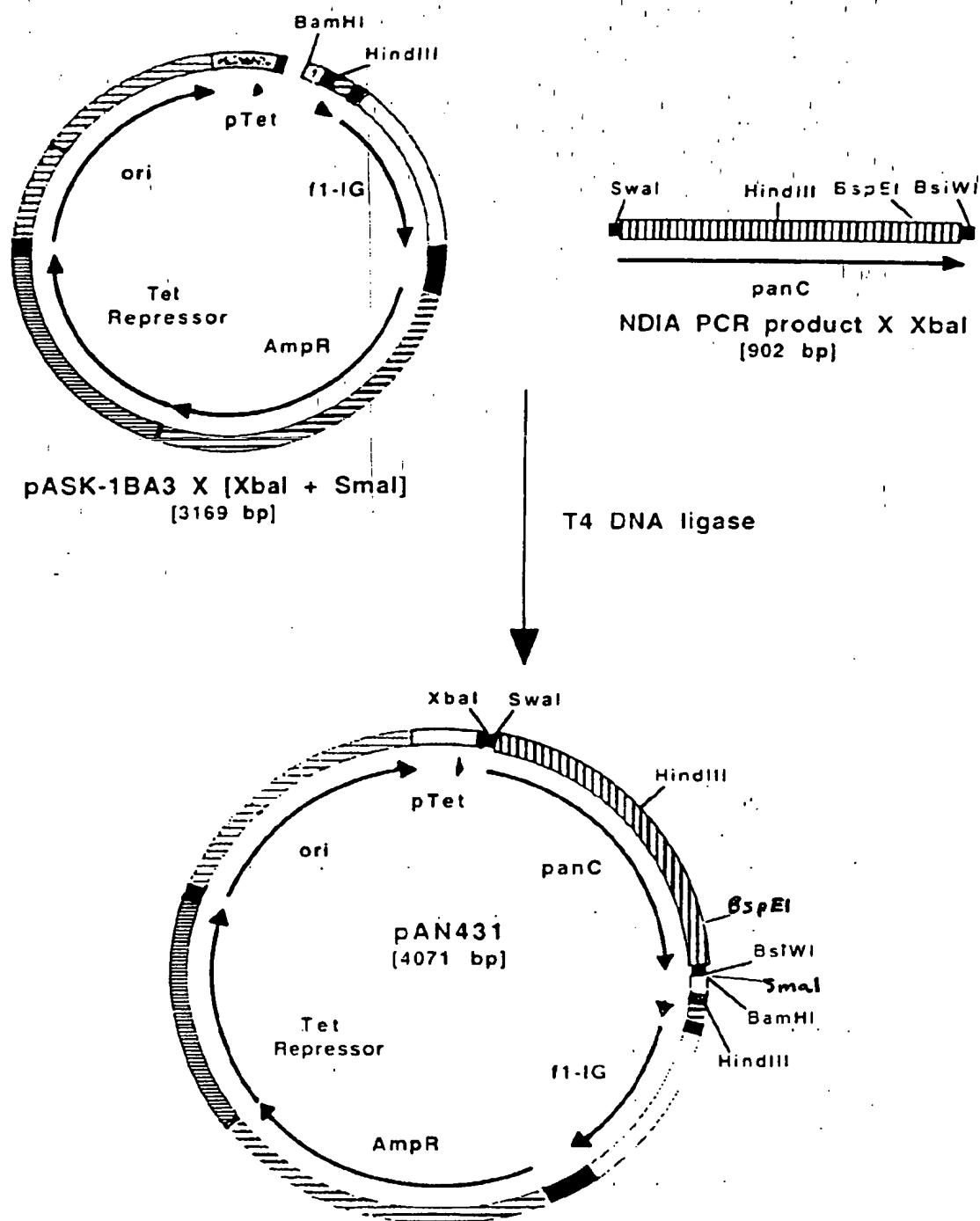
*Figure 8. Construction of pAN431.*

Figure 9. Construction of pAN441.

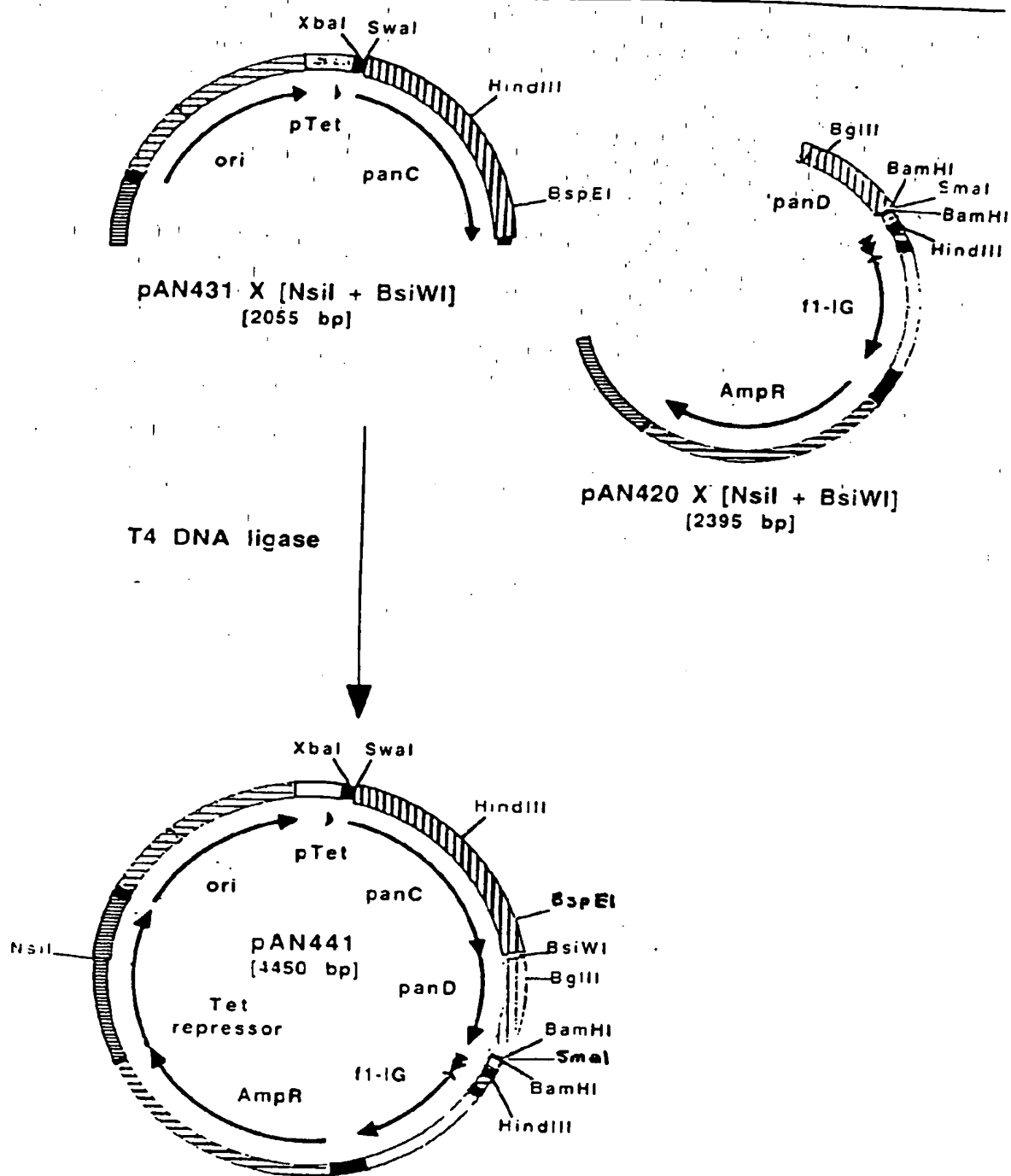
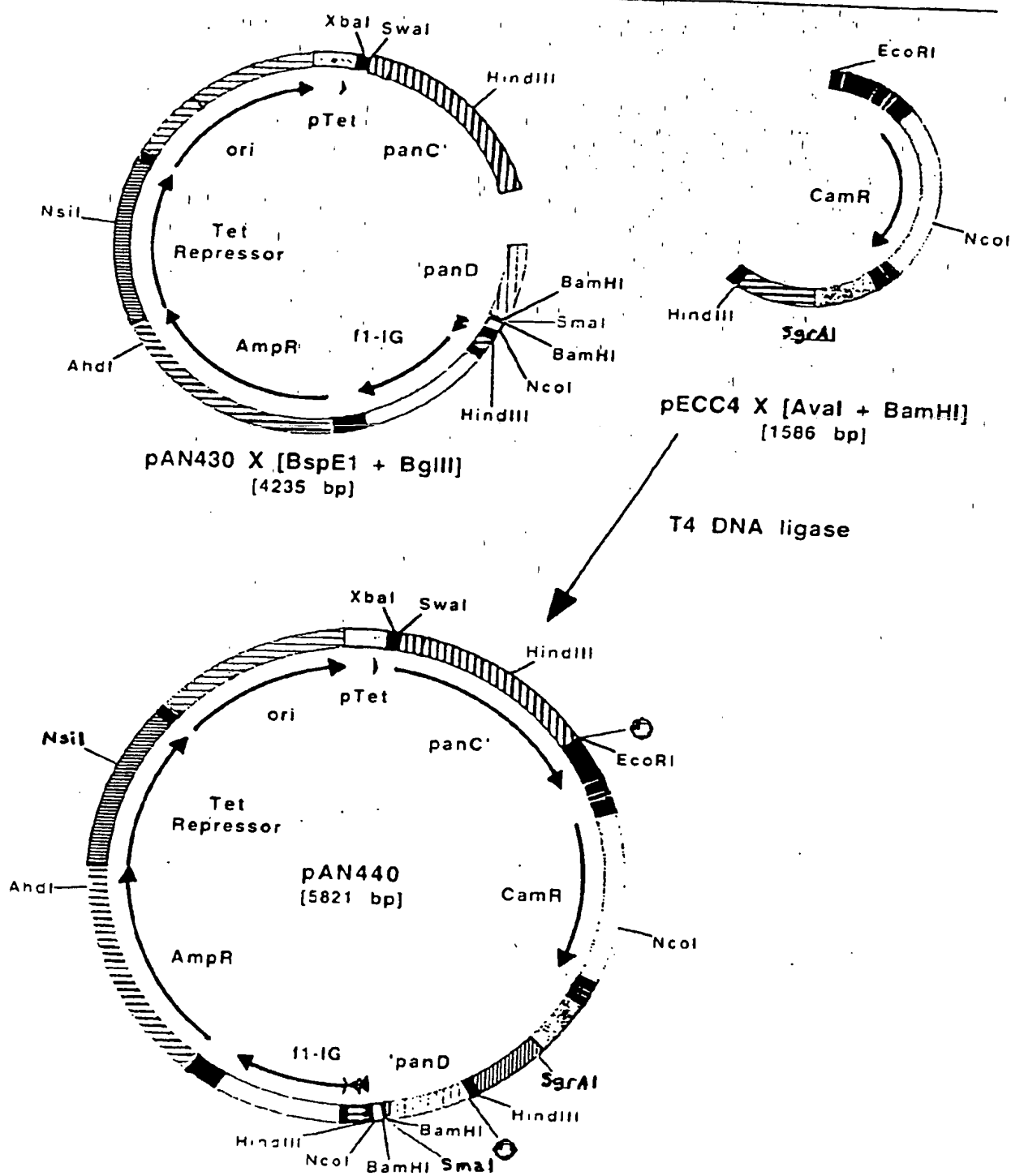
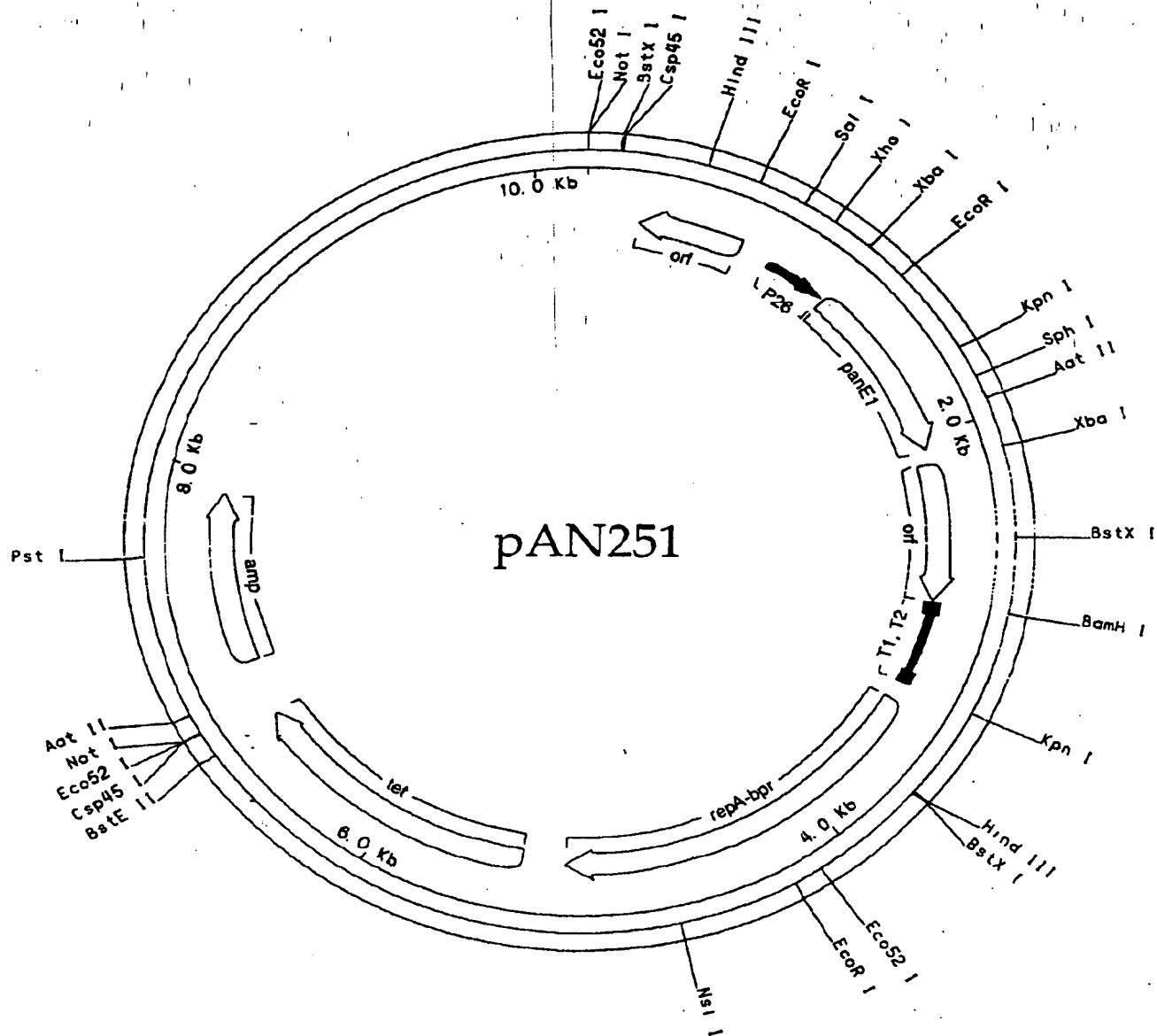


Figure 10. Construction of pAN440.



12/31

Figure 1| Structure of pAN251, a plasmid designed to integrate a single copy of P<sub>26</sub> panE1 at the panE1 locus by double crossover.



NSDOCID: &lt;WO\_\_0121772A2\_I\_&gt;

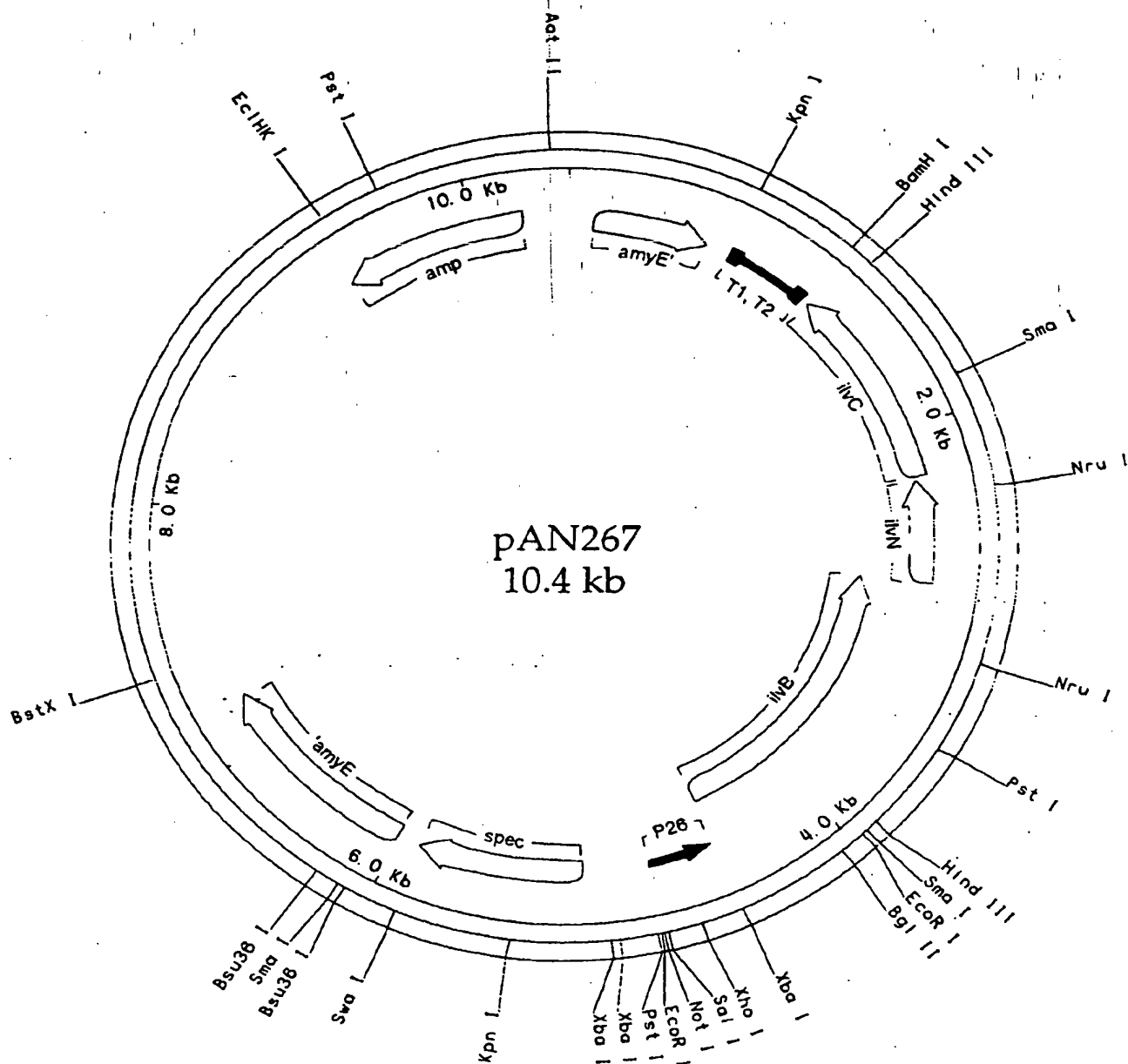
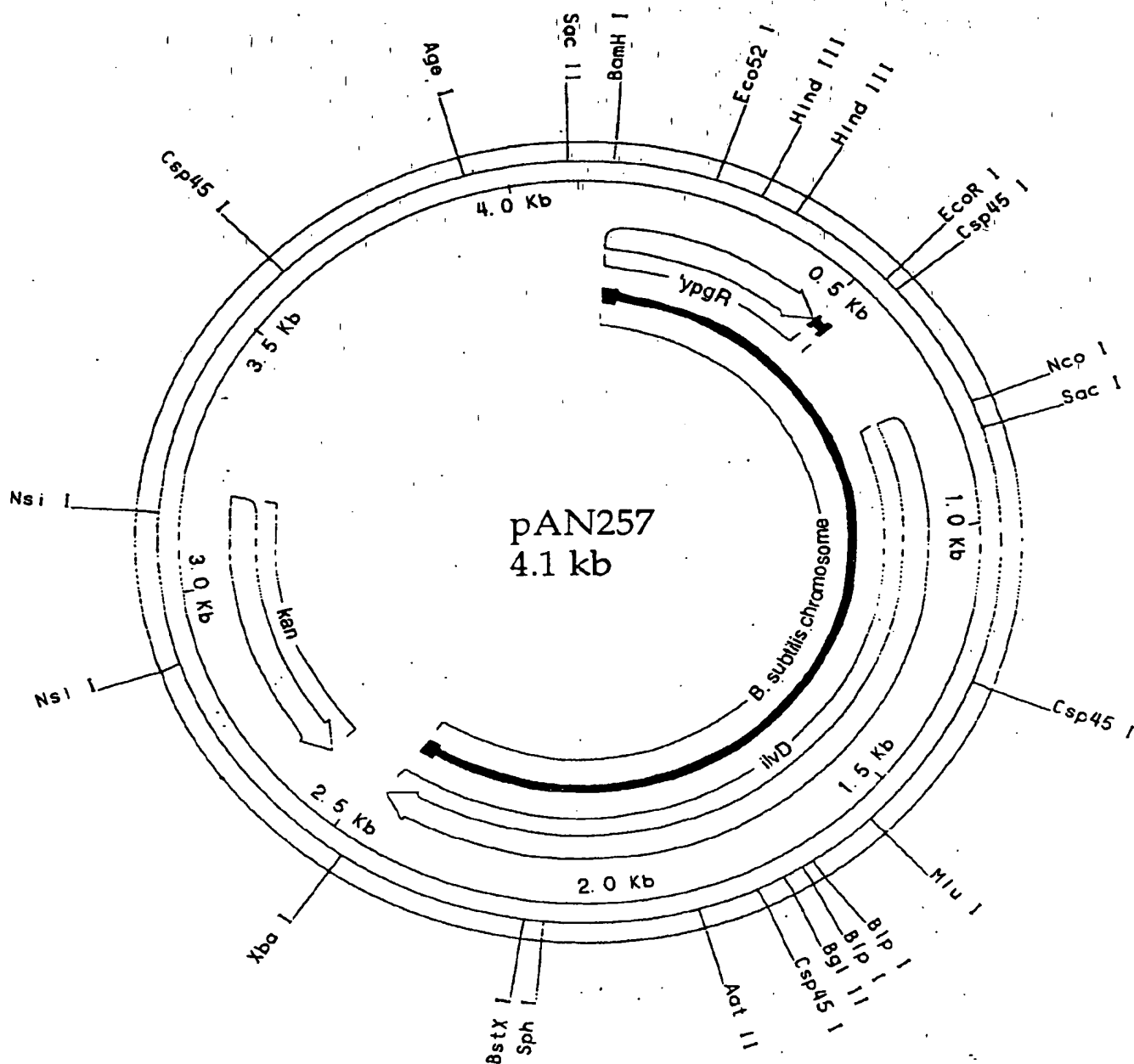


Figure 13 Structure of pAN257, a clone of *B. subtilis* *ilvD* in a low copy vector.



**Figure 14** Structure of pAN263, designed to stably integrate a single copy of P<sub>26</sub> ilvD at the ilvD locus.

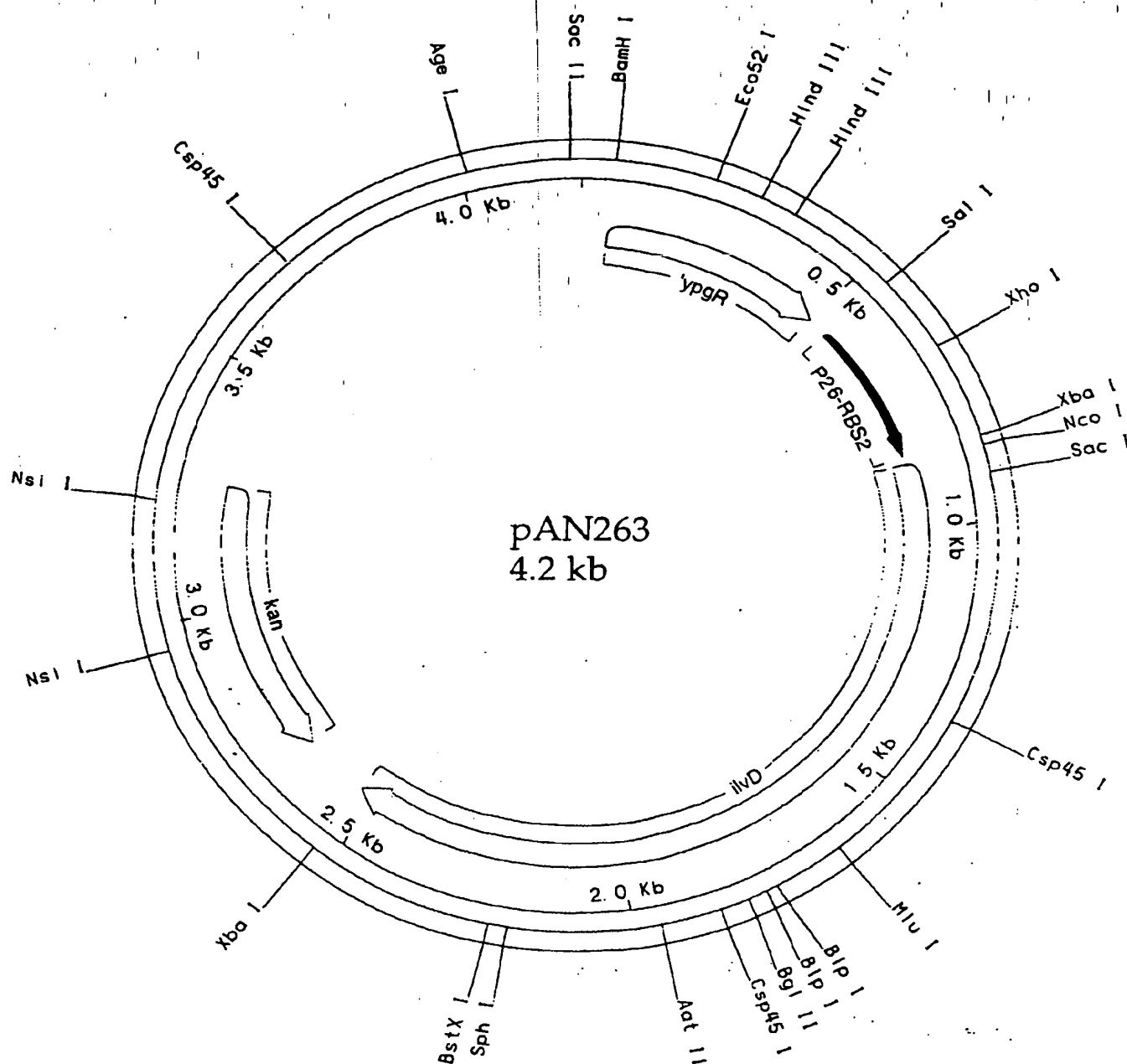


Figure 15 Structure of pAN261, designed to disrupt the *B. subtilis* *ilvD* gene with the *cat* gene.

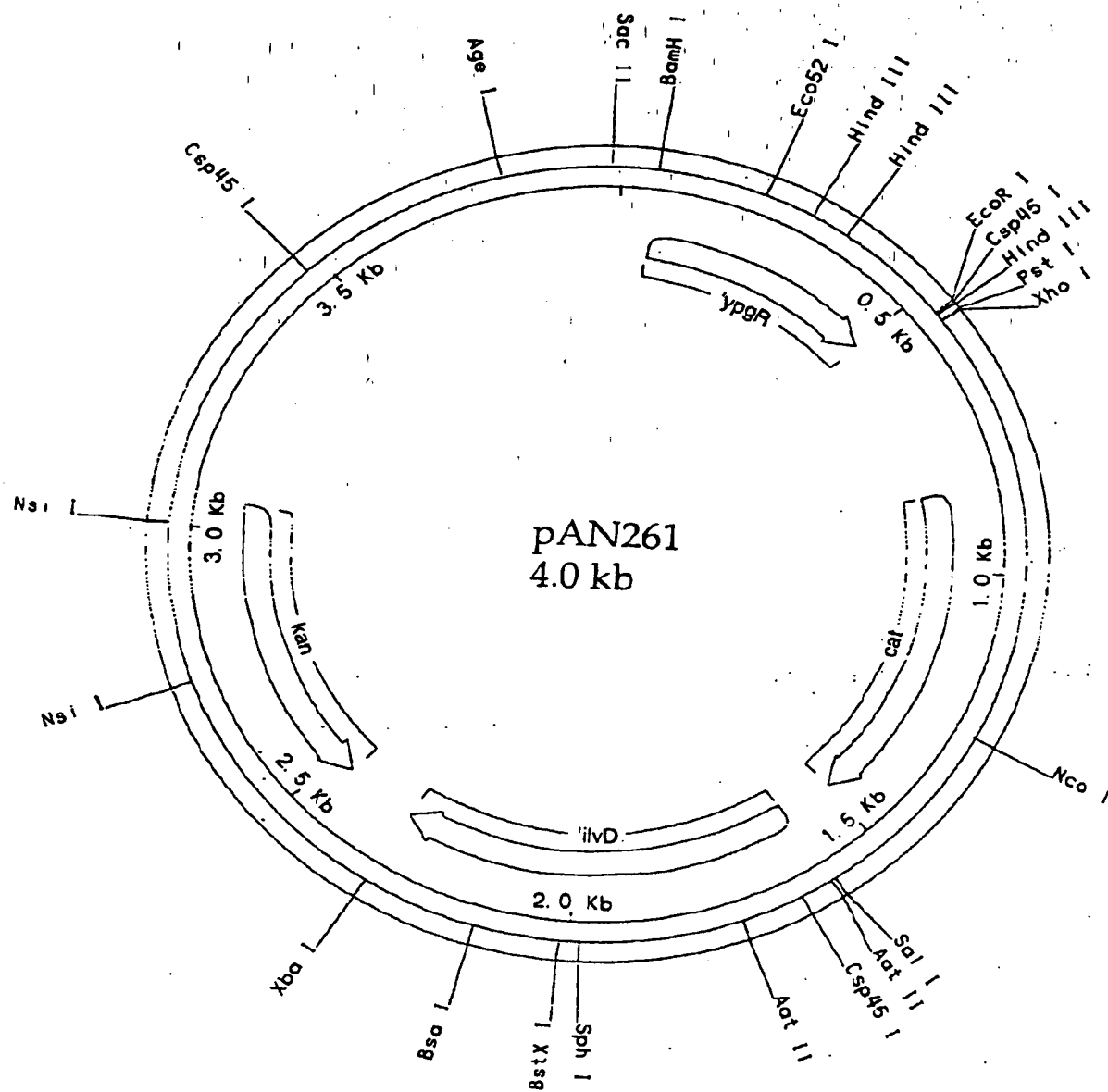
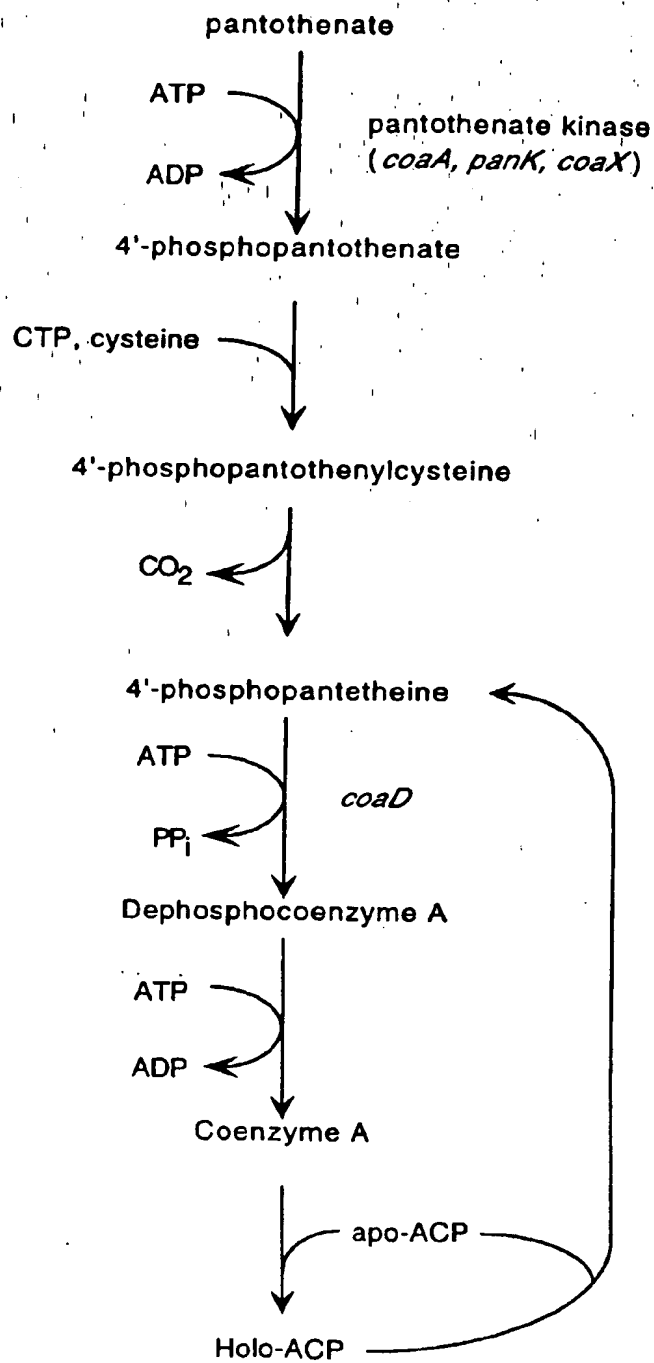
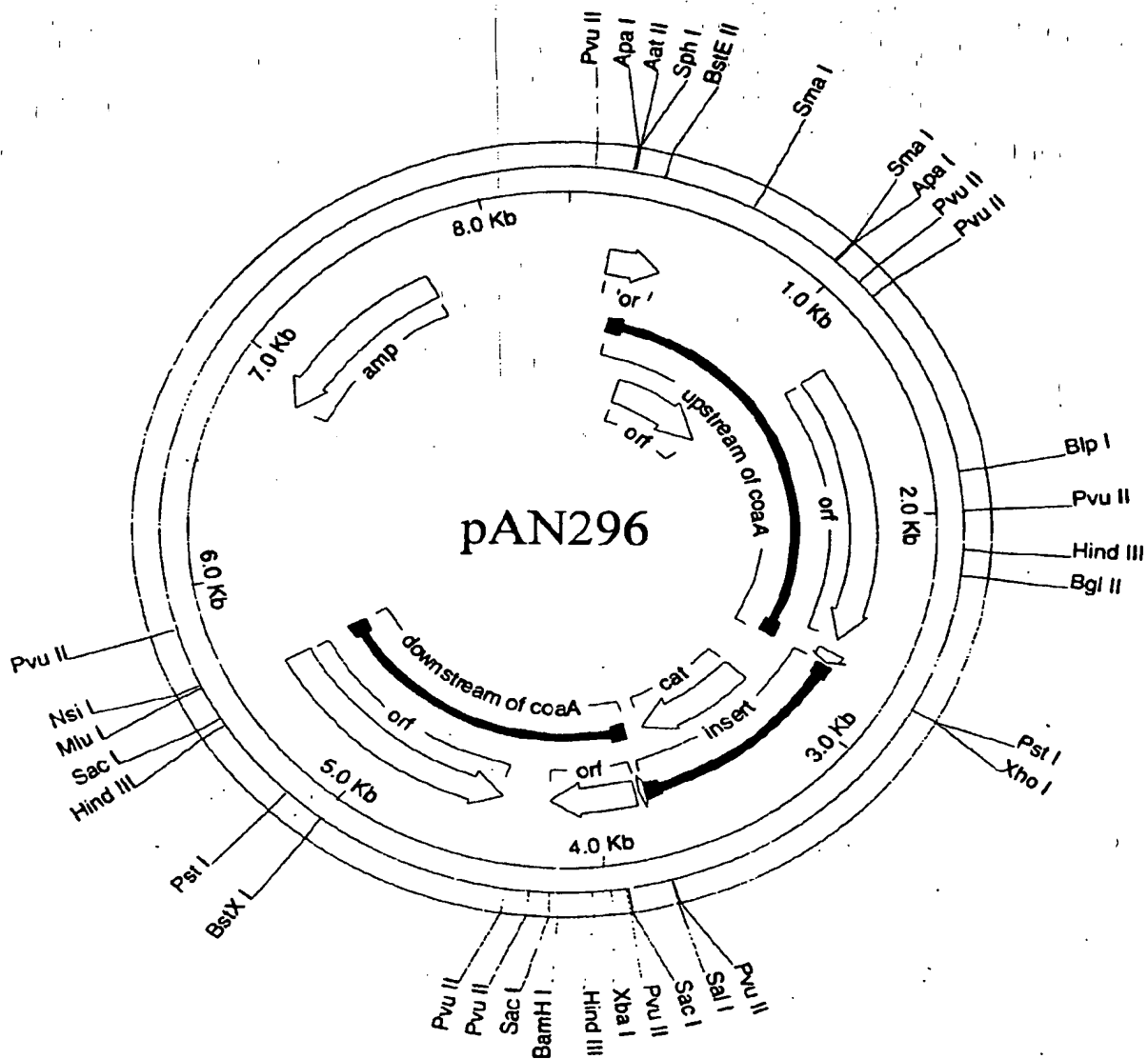




Figure 16

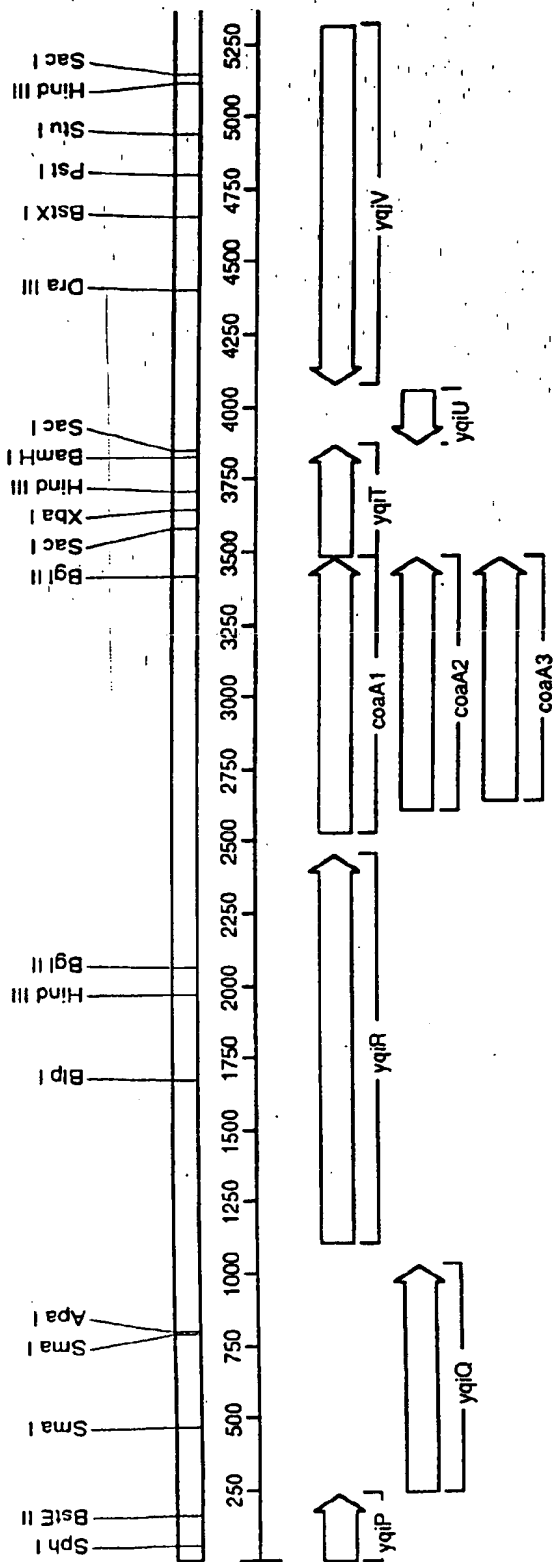


*Figure 7 Structure of pAN296, designed to delete most of the B. subtilis c<sub>o</sub>A gene and substitute a chloramphenicol resistance gene.*

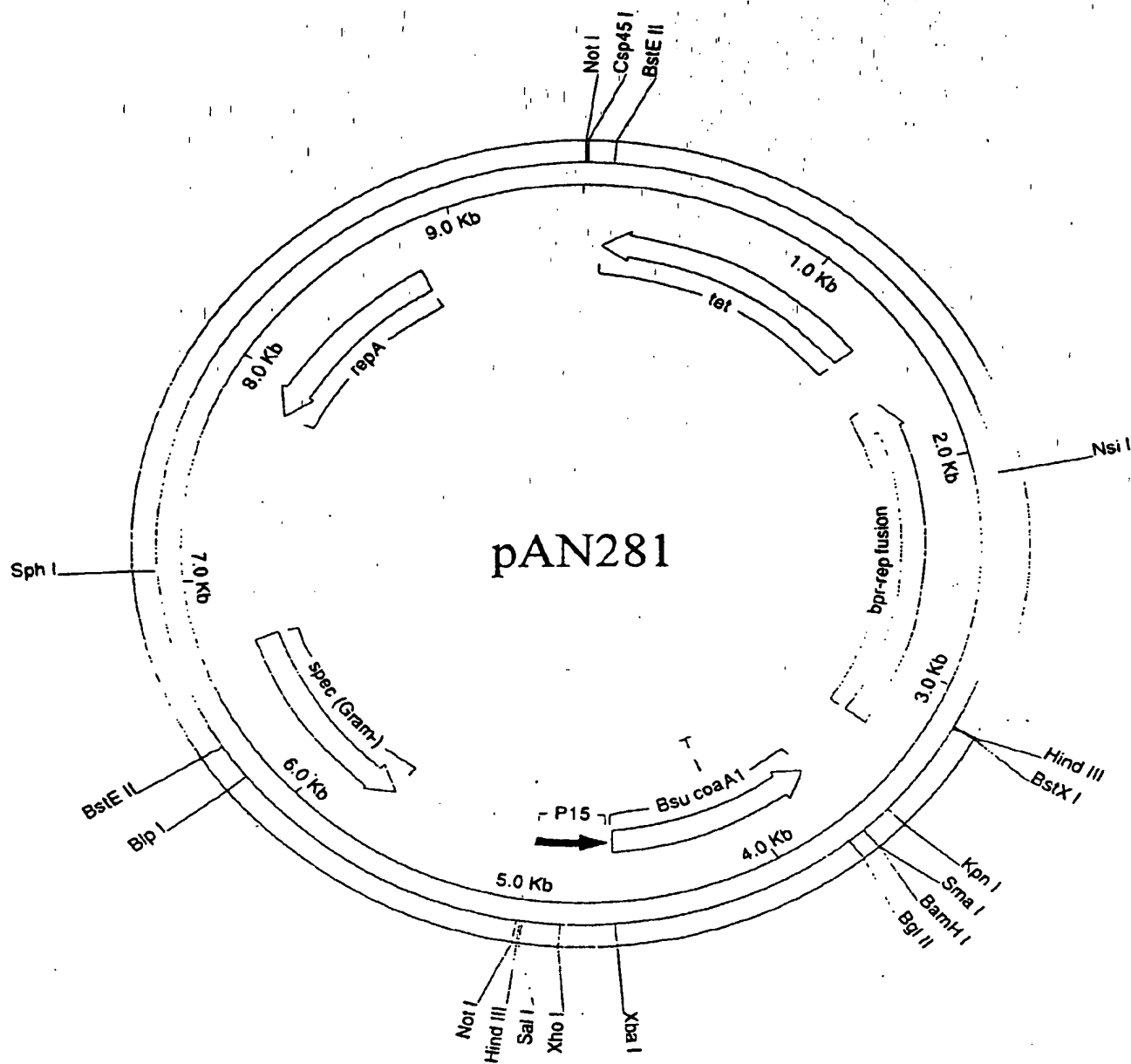


19/31

**Figure 18** Structure of the *B. subtilis* chromosome in the region of the *coaA* gene. The scale is in base pairs and the significant open reading frames are shown by the open arrows.



**Figure 19** Structure of pAN281, a plasmid for expressing *B. subtilis* *coaA* after integration at the *bpr* locus. pAN282 and pAN283 have similar structures.



### CLUSTAL W (1.7) Multiple Sequence Alignments

Sequence type explicitly set to Protein  
Sequence format is Pearson

Sequence 1:	spi Q9X795 M.leprae	312 aa
Sequence 2:	spi O86779 S.coelicolor	329 aa
Sequence 3:	spi O53440 M.tuberculosis	312 aa
Sequence 4:	spi P54556 B.subtilis	319 aa
Sequence 5:	spi P44793 H.influenzae	311 aa
Sequence 6:	spi P15044 E.coli	316 aa

sp Q9X795 M. leprae	-----P-----SPYVEFDRKQWRALRMSTPLATTEELIGLR
sp O53440 M. tuberculosis	-----MSRLSE-----P-----SPYVEFDRRQWRALRMSTPLATTEELVGLR
sp O86779 S. coelicolor	MISVPESIPRSRAHRQRPEATPYVDLTPREQWALRDKTPLPLTAAEEVEKLR
sp P44793 H. influenzae	-----MEFSQT-----QTPLTFNREQWALRKSVPKLATTEQDLKLL
sp P15044 E. coli	-----MSIKEQTL-----MTPYLQFDNRQWALRDSVPMTSLSEDEIARLK
sp P54556 B. subtilis	-----MKNKELN-----LHTLYTQHNRESWSGGFGLHSIAVSEEEAKAVE

sp|Q9X795|M.leprae  
sp|O53440|M.tuberculosis  
sp|O86779|S.coelicolor  
sp|P4793|H.influenzae  
sp|P15044|E.coli  
sp|P54556|B.subtilis

GLGEQIDLLEVEEYVPLRLRIHLQVAARQRLFAATAEFLGEFQOQNPGR  
 GLGEQIDLLEVEEYVPLRLRIHLQVAARQRLFAATAEFLGEFQOQNPDR  
 GLGQVIDLDEVROIYVPLSLRLNLVYGATDGLRGALNTFLGE--QGSG-  
 GNFEDLSLDEVSTIYVPLTRLINYYIDENLHRQTVLHREFLGR---NNAK  
 GNFEDLSLDEEVAETIYVPLRLNLNFYISNLRQAVLEQFLGT----NGOR  
 GLNDLYSLVEEETIYVPLRLNLHLHVKSAAERNKHVNVLKHPI---HSAK  
 \* . . . . \* \* \* \* \* \* \* \* \* \* \* \*

sp|Q9X795|M.leprae  
sp|O53440|M.tuberculosis  
sp|O86779|S.coelicolor  
sp|P44793|H.influenzae  
sp|P15044|E.coli  
sp|P54556|B.subtilis

[illegible]

FIG.20B

sp Q9X795 M.leprae	GRRLMHRKGFPESSYNRRALMRFTSVKSGADYACAPVYSHLRYDTIPGA
sp O53440 M.tuberculosis	QRRNLMRKGFPESSYNRRALMRFTSVKSGSDYACAPVYSHLHYDIIPGA
sp O86779 S.coelicolor	EARGLSRKGFPESSYRRALTRFVADIKAGKAEVTAPVYSHLIYDIVPDQ
sp P44793 H.influenzae	KQNLLOKKGFPVSYDTPKLIKFLADYKSGKSNVTAPIYSHLTYDIIPDK
sp P15044 E.coli	KERGLMKKGFPESSYDMHRLVRFVSDLKSGVPNVTAPVYSHLIYDVIPOG
sp P54556 B.subtilis	KKKNMSRKGFPESSYDVKALLEFLNDLKSGKDSVKAPVYSHLTYDREEGV
	::: : **** ** : * : : * : * : * : * : * : * : * : *
sp Q9X795 M.leprae	KHVVRHPDILILEGLNVLTGP-----TLMVSDLFDFSLYVDARIQD
sp O53440 M.tuberculosis	EQVVRHPDILILEGLNVLTGP-----TLMVSDLFDFSLYVDARIED
sp O86779 S.coelicolor	RLVVRHPDILILEGLNVLPALPGKGRGRT-RVGLADYDFDSVYVDARTED
sp P44793 H.influenzae	EDVVDKPDILILEGLNVLTGNK---TD-QTEVSDFVDFSIYVDAEKL
sp P15044 E.coli	DKTVVQPDILILEGLNVLSGMDYPHDPH-HVFVSDFVDFSIYVDAPEDL
sp P54556 B.subtilis	FEVVEQADIVIEGINVLSPTIEDDRENPRI FVSDFFDFSIYVDAEESR
	* : : * : : * : : * : : * : : * : : * : : * : : * : : * : : *
sp Q9X795 M.leprae	IEQWYVSRFLAMRGTAFA DPESHFHYSALTDSKAI IAAREIWR SINRPN
sp O53440 M.tuberculosis	IEQWYVSRFLAMRTTAFADPESHFHYYA FSDSQAVVAAREIWR TINRPN
sp O86779 S.coelicolor	IERWYLNRFKRLRATAFQNPSSYFRKYTQVSEEEALDYARTTWR TINRPN
sp P44793 H.influenzae	LKEWYIKRFLKRESAFNDPNSYFKHYASLSKEEAIATASKIWEINGLN
sp P15044 E.coli	LQTWYINRFLKREGAFTDPDSYFHNYSKLTKEEAIKTMTLWKEINWLN
sp P54556 B.subtilis	IFTWYLERFRLRETAFQNPDSYFHKFDLSQDEADEMAASIWE SYNRPN
	: * : : * : : * : : * : : * : : * : : * : : * : : * : : *
sp Q9X795 M.leprae	LVENILPTRPRATLVLRKADHSHINRLRLRKL
sp O53440 M.tuberculosis	LVENILPTRPRATLVLRKADHSHINRLRLRKL
sp O86779 S.coelicolor	LVENVAPTRGRATLVLRKGPDPHKVQRLSLRKL
sp P44793 H.influenzae	LNQNILPTRERANLILKKGHNHQUELILKLRK-
sp P15044 E.coli	LKQNILPTRERASLILTKSANHAAVEEVLRLK-
sp P54556 B.subtilis	LYENILPTFRSDLILAKGDGHKVEEVLVRRV
	* : : * : : * : : * : : * : : * : : * : : * : : * : : *

Figure 21

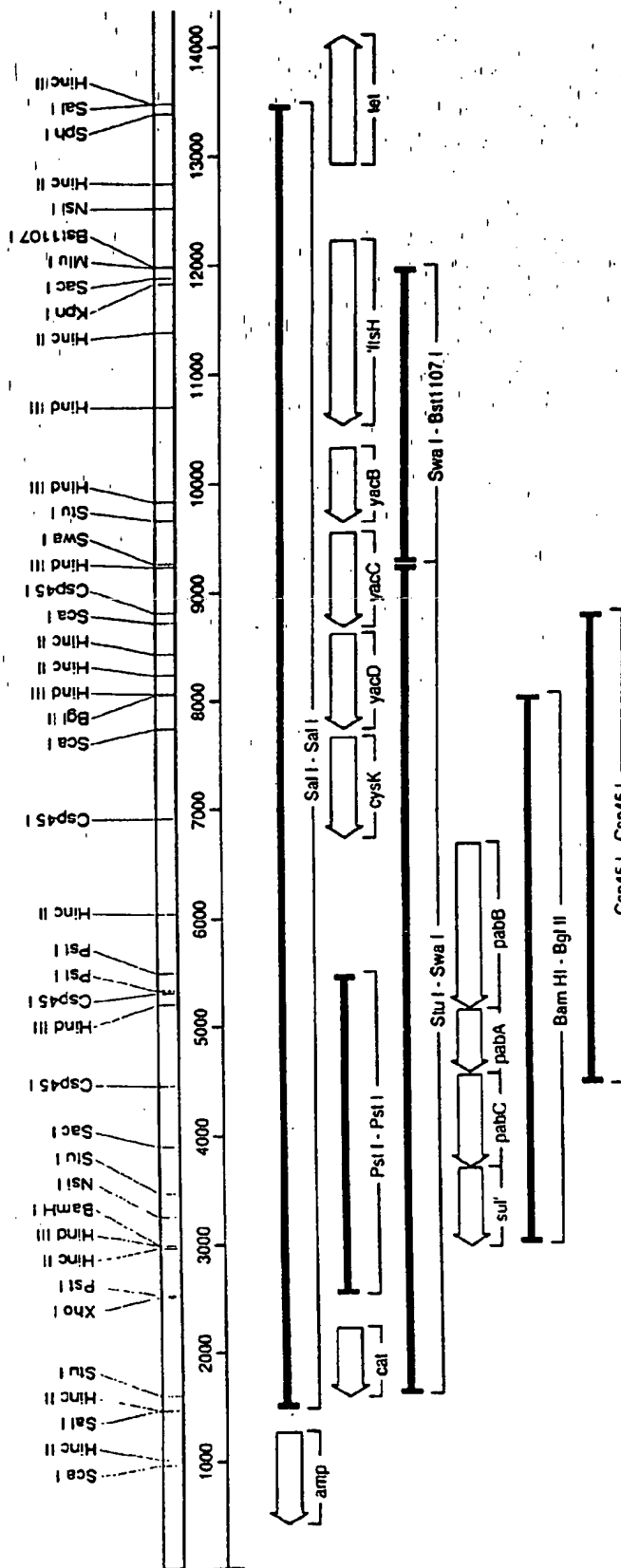
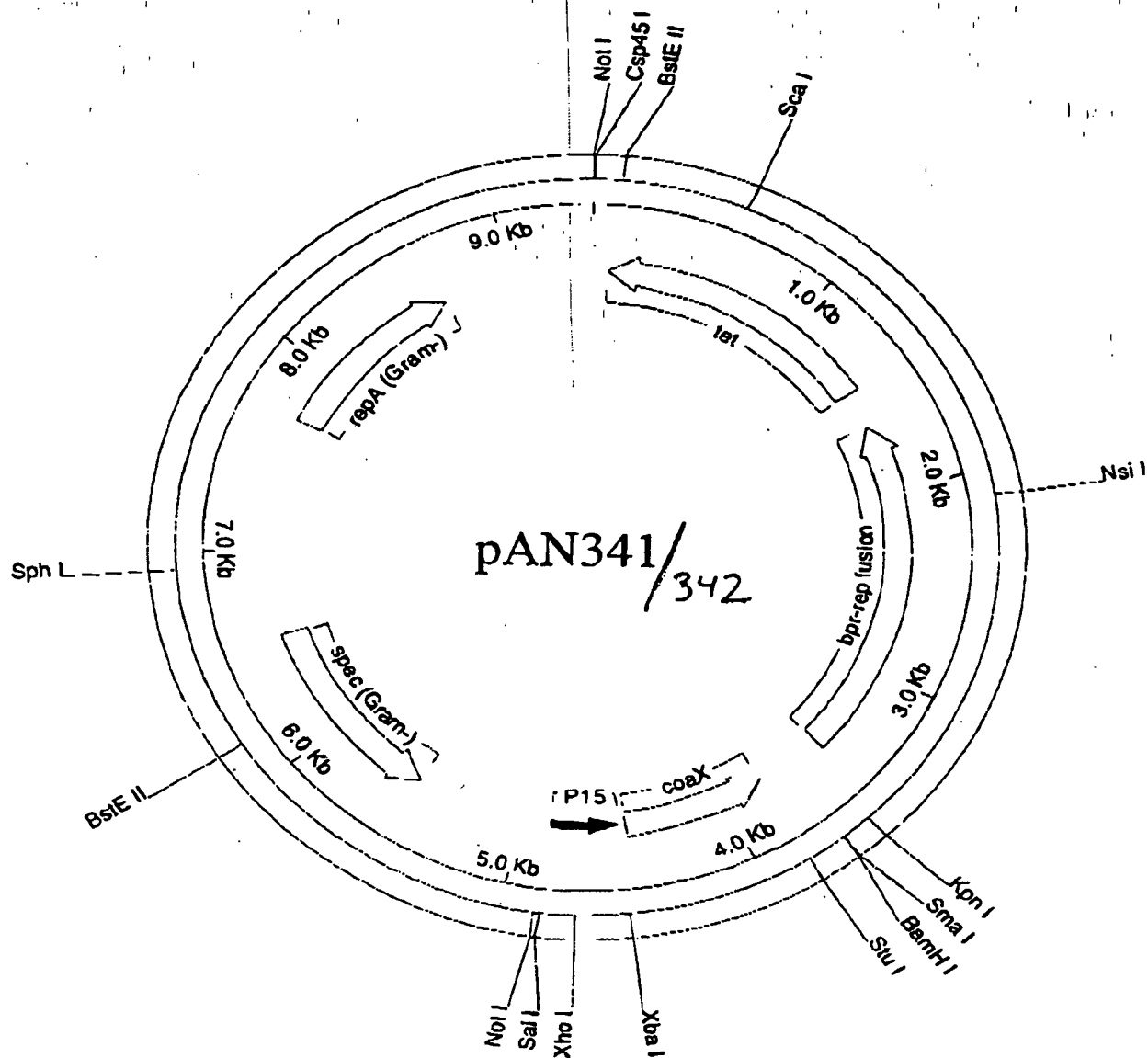


Figure 22 Structure of pAN341 and pAN342, two independent PCR-derived clones of *yacB* (renamed *coaX*).





## FIG.23A

## CLUSTAL W (1.7) Multiple Sequence Alignments

Sequence type explicitly set to Protein

Sequence format is Pearson

Seq. 1: B.subtilis|Coax|SEQNO\_9 258 aa Seq. 8: sp|O51477|B.burgdorferi 262 aa  
 Seq. 2: dbj|BAA21476.1|D.vulgaris 212 aa Seq. 9: sp|P74045|Synechocystis 257 aa  
 Seq. 3: gb|AAD35964.1|T.maritima 246 aa Seq. 10: sp|O25533|H.pylori 223 aa  
 Seq. 4: pir|T36391|S.coelicolor 265 aa Seq. 11: sp|O67753|A.aeolicus 229 aa  
 Seq. 5: sp|Q45338|B.pertussis 267 aa Seq. 12: sp|Q9RX54|D.radiodurans 262 aa  
 Seq. 6: sp|O06282|M.tuberculosis 272 aa Seq. 13: WIT|RCA03301|C.acetobutylicum 250 aa  
 Seq. 7: sp|O83446|T.pallidum 273 aa Seq. 14: WIT|RC02473|R.capsulatus 258 aa

B.subtilis|Coax|SEQIDNO\_9 -----MLLVIOVGNTNTVLGVYHDG-----KLEYHWRIE  
 WIT|RCA03301|C.acetobutylicum NKRAAFMLLLFLRSVLKVLVLDVGNTNIVGIYNDT-----KLTAEWRLS  
 pir|T36391|S.coelicolor -----MLLTIDVGNTHTVLGLFDGE-----DIVEHWRIS  
 sp|O06282|M.tuberculosis -----MLLAIDVRNTHTVVGLSGMKEHAKVQWQRIR  
 WIT|RC02473|R.capsulatus -----MLLCIDCGNTTVFSVWDGT-----DFAATWRIA  
 dbj|BAA21476.1|D.vulgaris -----MTQHFLLEDIGNTNKIGIAVET-----AVLTSVLP  
 sp|Q9RX54|D.radiodurans -----MPAPPLLAVIDIGNTTVLGLADASG-----ALHTTWIR  
 gb|AAD35964.1|T.maritima -----MYLLVDVGNTHSVFSITEDG-----KTFRRWRLS  
 sp|O83446|T.pallidum -----MLLIDVGNSHVVFIOGENGGRVCVRELFLA  
 sp|O51477|B.burgdorferi -----MNKPLLSELIIDIGNTSIAFALFKDN-----QVNLFKMK  
 sp|O67753|A.aeolicus -----MRFLTVDVGNSVDIALWEGK-----KVK  
 sp|P74045|Synechocystis -----METSKECCGLALDNDKQKPWLGLMIGN-----SRLHWAYC  
 sp|O25533|H.pylori -----MPARQSFDTLKN-----LVLCDIGN-----TR  
 sp|Q45338|B.pertussis -----MIILIDSGNSRLKVGWFDPDAP--QAAREAPV

FIG.23B

TSRHKTEDEFGMILRSFDHS---GLMFEQIDGIIISVVPPIMFALER  
 TDVLSADEYGIQVNLFOQD---KLDPTLVEGVIISSVVPNIMYSLEH  
 TDSRRRTADELAVLQGLMGHPLLGDELGDGIDGIAICATVPSVLHELRE  
 TESEVTADELALITDGLIG-----EDSERLTGTAALSTVPSVLHEVRI  
 TDHRRRTADEYFVWLNTLMQLK---GLQGRISEAIISSAPRVVFNLRV  
 TDPGQTTDSIGRLLEVLRHAG---LGPADVGACVASSVPGVNPILRR  
 TNREMLPDDLALQLHGLFTLA---GAP-IPRAAVLSSVAPPVGENYAL  
 TGVFQTEDELEFSLHPLLG-----DAMREIKGIGVASVVPQTNTVIER  
 PDARKTQDEYSLLLHALCERAG---VGRASLRDAFISSVVPVLTKTIAO  
 TNLMRLRYDEVYSFTEENFDEN-----VN---K-VFISVVPIILNETFKN  
 DELKLSHEEFLEKEFPKLK-----ALGISVKQSFSEKVRG  
 SGNAPLQWTVDYNPKSAQLP-----VLLGKVPMLLASVWPE  
 IHFAQNYQLFSSAKEDLKR-----LGIQKEIFYISWNEE  
 AFDNLDLDALGRWLATLPRRP-----Q-----RALGVNVAGLARGEAIA

MCTKYFHIEPQIVG-PG-MKTGLNICKYDNPKVEVGADRIVNAAVAIHLYG-  
 MIRKYFKINPLVVG-PG-IKTGINICKYDNPKVEVGADRIVNAAVAHEIYK-  
 VTRRYGQDPAVLVEPG-VKTGVPIILTDHPKEVGADRIINAAVAVELYG-  
 MLDQYWPSVPHVLIIEPG-VRTGIPLLVDNPKVEVGADRIVNCCLAAVDRFR-  
 LCNRYFDCRPYVVGKPG-CELPVAPRVDPGTTVGPDLVNTVAGVDRHG-  
 ACERYL--YRKLFPAGDIAIPLDNRYERPAEVGADRLVAAYAAARLLYP-  
 ALKRHEMIDAFVSAEN--LPDVTVELDTPGSVGADRLCNLFCAEKYLG-  
 FSQKYEHISPIWVKAKN---GCVKNNVKNPSEVGADRVANVAVFYKEYG-  
 AVAQISGVQPVVFGPWAYEHLVPVRIPEPVRAEIGTDLVANAAVAAYVHER-  
 VIFSEFKIKPLFIFGDLNVDLTENPYKSDKELGSDVFANLVAATENYS-  
 KIPKIK-----FLKEN---FPIQVDYKTPETLGTDRVALAYSARKFYG-  
 QTEVWRVQPKILTILKN---LPLVNLYP---SFGIDRALAGLGTGLTYG-  
 NEKALLNCYPNAKNTAG--FFHLETDYVG---LGIDRQVACLA---VN--  
 ATL RAGGCDIRWLRAQP-LAMGLRNGYRNPDQIGADRWACMVGVLARQPS

B. subtilis|Coax|SEQIDNO\_9  
 WIT|RCA03301|C.acetobutylicum  
 pir|T36391|S.coelicolor  
 sp|O06282|M.tuberculosis  
 WIT|IRC02473|R.capsulatus  
 dbj|BAA21476.1|D.vulgaris  
 sp|Q9RX54|D.radiodurans  
 gb|AAD35964.1|T.maritima  
 sp|O83446|T.pallidum  
 sp|O51477|B.burgdorferi  
 sp|O67753|A.aeolicus  
 sp|P74045|Synecocystis  
 sp|O25533|H.pylori  
 sp|Q45338|B.pertussis

B. subtilis|Coax|SEQIDNO\_9  
 WIT|RCA03301|C.acetobutylicum  
 pir|T36391|S.coelicolor  
 sp|O06282|M.tuberculosis  
 WIT|IRC02473|R.capsulatus  
 dbj|BAA21476.1|D.vulgaris  
 sp|Q9RX54|D.radiodurans  
 gb|AAD35964.1|T.maritima  
 sp|O83446|T.pallidum  
 sp|O51477|B.burgdorferi  
 sp|O67753|A.aeolicus  
 sp|P74045|Synecocystis  
 sp|O25533|H.pylori  
 sp|Q45338|B.pertussis

FIG.23C

B. subtilis Coax SEQIDNO_9	NP--LIWVDFGATTYCYIDENKQVMGGAIPGITISTEALYSRAAKLPR	IEITRPDN---IIGKNTVSAMQSGILFGYVGVQVEGIVKRMKWAQKDLK-
WITIRCA03301 C.acetobutylicum	RS--LIIDFGTATTFCAVRENGDYLGGAICPGIKVSSSEALFEKAALPR	VELIKPAY---AICKNNTISSIQSGIVYRVRQVKYLFKELKENLPDGRRT
pir T36391 S.coelicolor	GP--AIWVDFGATTTFDAVSARGEYIGGVIAPGIEISVEALGVKGAQLRK	IEVARPRS---VIGKNTVEAMQSGIVYGFAGQVDGVVNRMARLADD--P
spi006282 M.tuberculosis	KA--AIWVDFGSSICVDVVSAGKEFLGGAIPGVQVSSDAAAARSAAALRR	VELARPRS---VVGKNTVECMQAGAVFGFAGLVGLVGRIRREDVSGFSVD
WITIRRC02473 R.capsulatus	GD--LIWVDFGATTFDVWAPDGAIVGGVIAPGVNLSLEALHMAAALPH	VDVTKPGQ---VIGTNTVACIQSGVYWGVI GLVEGIVRQIRMERDRP---
dbj BAA21476.1 D.vulgaris	GPRSLVDFGATTTFDCVEG-GAYLGGILCPGVLSAGALSSTRTAKLPR	ISLEVEEDS-PVIGRSTTSLNHGFI FGFAAMTEGVLA-------
spi09RX54 D.radiodurans	GLDYAVVDFGTFSTNFDVGRGRRELGGILATCAQVSADALFARAALPR	ITLQAPET---AIGKNTVHALQSGLVFGYAEVMDGLLRIRRAELPGE---
gb AAD35964.1 T.maritima	KN--GIIIDMGATTVDLVVN-GSYEGGAILPGFFMVVHSLFRGTAKLPL	VEVKPADF---VVGKDTENIRLGVVNGSVYALEGITIGRIKEVYGOLP---
spi083446 T.pallidum	SA--CVVDCGTALTFTAVDGTGLIQQVAIAPGLRTAVQSLHTGTQAQLPL	VPLALPDS---VLGKDTTHAVQAGVVRGTLEFVIRAMIAQCQKELGCR---
spi051477 B.burgdorferi	FEN-VLVVDLGTACTIFAVSRQDGLGGI INSGPLINFNSLDNDAYLIK	FPISSTNN---LLERTTSGVNSGLFYQYKYLIEGVYRDIKQYKKK---
spi067753 A.aeolicus	KN--VVVISAGTALVIDLVE-GKFKGGFTLGLGKKLKILSOLAEGIPE	FFPEEVEI---FLGRSTRECVLGGAYRESTEFIKSTLKLWRKVFKRK---
spiP74045 Synechocystis	FP--CLVVDGGTALTITGFDQDKLVGGAILPGLGLQLATLGBRLAALPK	LEMDQTELPRWALDTPSAIFSGVVYGVLGALQSYLDQDWQKLPFGA---
spi025533 H.pylori	NG---VVDAGSAITIDLIKE-GKHLGGCILPGLAQYI HAYKKSAKILEQ	PFKALDSL--EVLPKSTRDAVNYGMVLSVIACIQHLAK--NQK-----
spiQ45338 B.pertussis	VHPPLLVASFGTATTLDITGPONVFPFGLILPGPAMMRGALAYGTAHLPL	ADGLVADY-----PIDTHQAIASGIAAAQACAIVRQWLAGRQRYGAP--

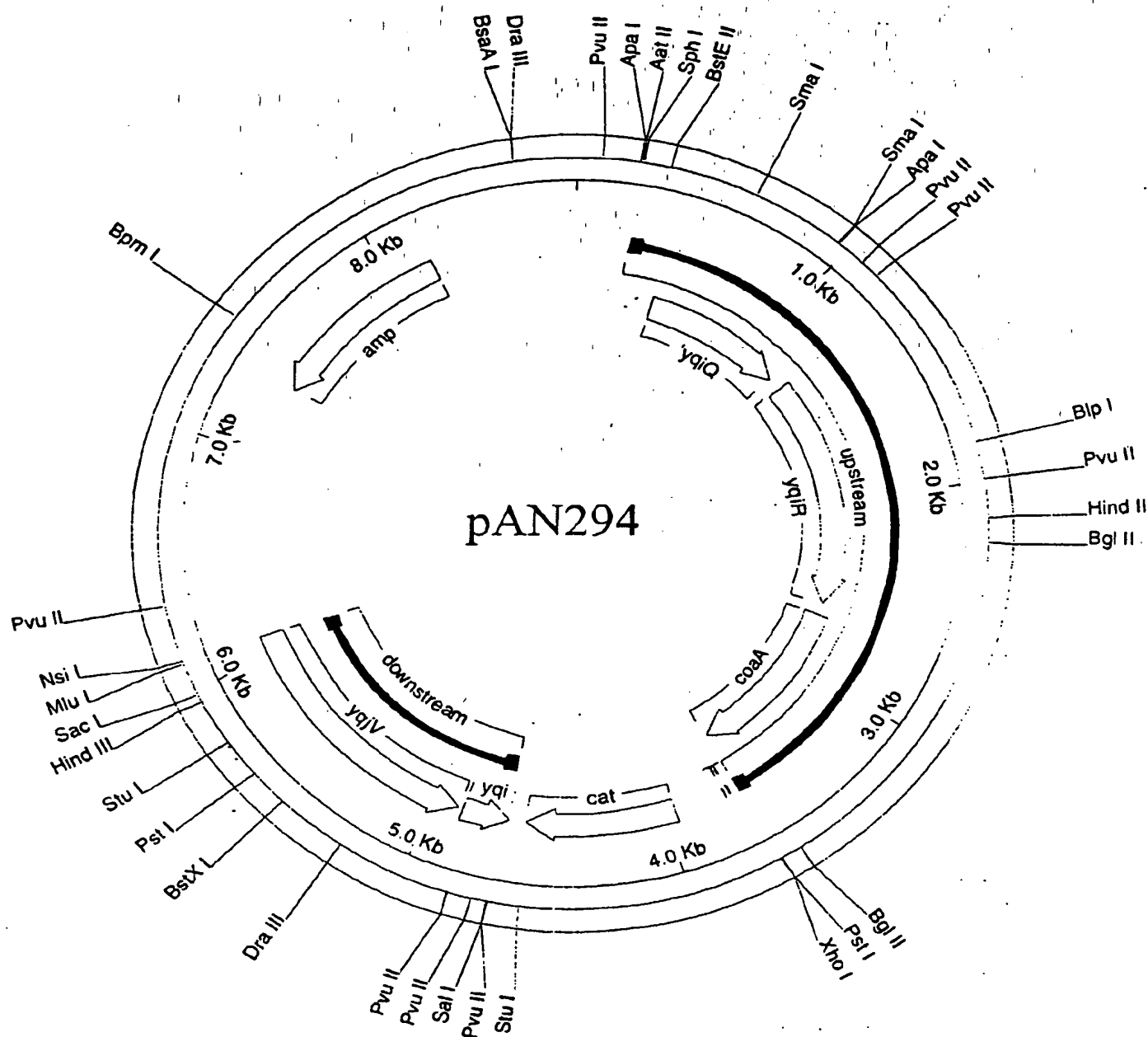
FIG.23D

B. subtilis Coax SEQIDNO_9	-----VIATGG-----LAPLIANES-----DCIDIVDPFLLKGLELI
WIT RCA03301 C.acetobutylicum	RTSLVLTGG-----LAKLIN-----
pir T36391 S.coelicolor	DDVTVTATGG-----LAPVLGES-----SVIDEHEPWLTLMGLRLV
spi O06282 M.tuberculosis	HDVAIVATGH-----TAPLLPEL-----HTVDHYDQHLTLOGLRLV
WIT IRC02473 R.capsulatus	--MKVIATGG-----LASLFDLGF-----DLFDKVEDDLTMHGLRLI
dbj BAA21476.1 D.vulgaris	-----
spi Q9RX54 D.radiodurans	--AVAVATGG-----FSRTVQIC-----QEIOYDETLRLGLVEL
gb AAD35964.1 T.maritima	-----VVLTTG-----QSKIVK-DM-----IKHEIFDEDLTKGVYHF
spi O83446 T.pallidum	--CAAVITGG-----LSRLFS-SE-----VDFPPIDAQLTSLGLAHI
spi O51477 B.burgdorferi	--FNLIITGG-----NADLILSLI-----EIEFIFNIHLTVEGVRIL
spi O67753 A.aeolicus	--FKVITGG-----EGKYFS-----KFGIYDPLLVRHGMRL
spi P74045 Synechocystis	--AMVITGG-----DGKILHGFELKEHSPNLSVAWDDNLI FLGMAAI
spi O25533 H.pylori	-----IYLCGG-----DAKYLSAFL-----PHSVCKERLVFDGMEIA
spi Q45338 B.pertussis	---EIYVAGGGWPEVRQEAERLLAVTGAAFGATPQPTYLDSPVLDGLAAL
B. subtilis Coax SEQIDNO_9	YERNRVGSV-----
WIT RCA03301 C.acetobutylicum	-----
pir T36391 S.coelicolor	YERNVSRM-----
spi O06282 M.tuberculosis	FERNLEVQRGLKTAR-----
WIT IRC02473 R.capsulatus	FDYKGLGA-----
dbj BAA21476.1 D.vulgaris	-----
spi Q9RX54 D.radiodurans	WASRSEVR-----
gb AAD35964.1 T.maritima	CFGD-----
spi O83446 T.pallidum	ARLVPTSLPPATVSGSSGN
spi O51477 B.burgdorferi	GNSIDFKFVN-----
spi O67753 A.aeolicus	LYLYHRI-----
spi P74045 Synechocystis	HGDRPIC-----
spi O25533 H.pylori	LKKAGILECK-----
spi Q45338 B.pertussis	AAQGAPTA-----

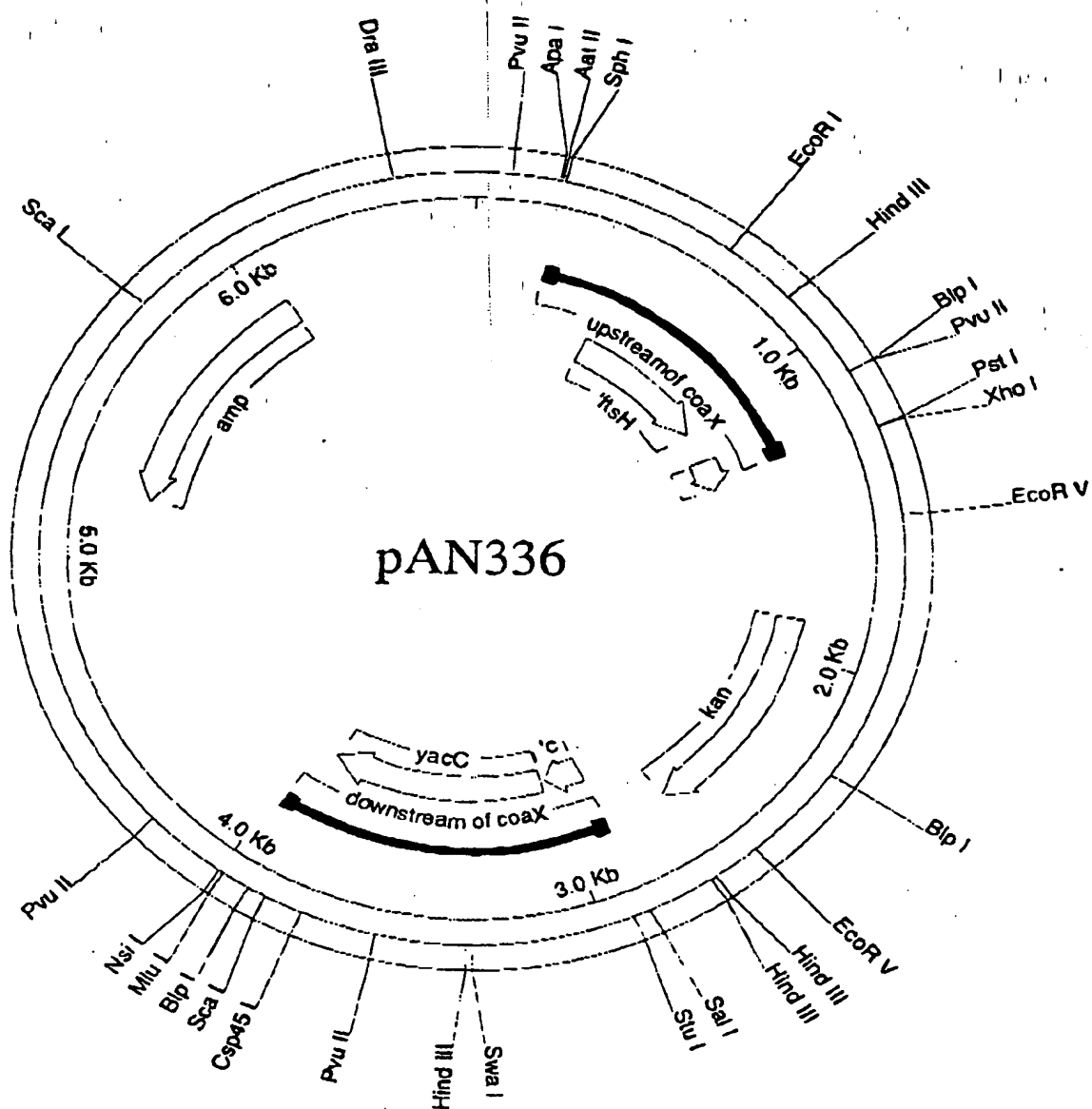
Figure 24 Alignment of a portion of the amino acid sequences of several known or suspected pantothinate kinases. The residues that are mutated in *E. coli* coaA15(Ts) and *B. subtilis* coaA from plasmid pAN282A are indicated below and above the alignment, respectively. The coordinate given in the left margin for the *B. subtilis* protein refers to the coaA1 open reading frame.

	K	D	N	V	T	A	P	V	Y	S	H	L	I	Y	D	I	I	P	G	A	Majority
168	K	D	S	V	K	A	P	V	Y	S	H	L	T	Y	D	R	E	E	G	V	<i>B. subtilis</i> CoaA1
167	V	P	N	V	T	A	P	V	Y	S	H	L	I	Y	D	V	I	P	D	G	<i>E. coli</i> CoaA
165	K	S	N	V	T	A	P	I	Y	S	H	L	T	Y	D	I	I	P	D	K	<i>H. influenzae</i> CoaA
169	A	D	Y	A	C	A	P	V	Y	S	H	L	R	Y	D	T	I	P	G	A	<i>M. leprae</i> CoaA
169	S	D	Y	A	C	A	P	V	Y	S	H	L	H	Y	D	I	I	P	G	A	<i>M. tuberculosis</i> CoaA
179	K	A	E	V	T	A	P	V	Y	S	H	L	I	Y	D	I	V	P	D	Q	<i>S. coelestis</i> CoaA

**Figure 25** Structure of pAN294, a plasmid for integrating mutagenized *B. subtilis* *coaA* at its native locus.



**Figure 26** Structure of pAN336, a plasmid designed to delete *B. subtilis* *coaX* from the chromosome and replace it with a kanamycin resistance gene.



- 1 -

## SEQUENCE LISTING

&lt;110&gt; OMNIGENE BIOPRODUCTS

<120> METHODS AND MICROORGANISMS FOR PRODUCTION OF  
PANTO-COMPOUNDS.

&lt;130&gt; BGI-141CPPC

&lt;140&gt;

&lt;141&gt;

&lt;150&gt; USSN 09/400,494

&lt;151&gt; 1999-09-21

&lt;150&gt; USSN 60/210,072

&lt;151&gt; 2000-06-07

&lt;150&gt; USSN 60/221,836

&lt;151&gt; 2000-07-28

&lt;150&gt; USSN 60/221,836

&lt;151&gt; 2000-08-24

&lt;160&gt; 94

&lt;170&gt; PatentIn Ver. 2.0.

&lt;210&gt; 1

&lt;211&gt; 311

&lt;212&gt; PRT

&lt;213&gt; Haemophilus influenzae

&lt;400&gt; 1

Met Glu Phe Ser Thr Gln Gln Thr Pro Phe Leu Ser Phe Asn Arg Glu  
1 5 10 15Gln Trp Ala Glu Leu Arg Lys Ser Val Pro Leu Lys Leu Thr Gln Gln  
20 25 30Asp Leu Lys Pro Leu Leu Gly Phe Asn Glu Asp Leu Ser Leu Asp Glu  
35 40 45Val Ser Thr Ile Tyr Leu Pro Leu Thr Arg Leu Ile Asn Tyr Tyr Ile  
50 55 60Asp Glu Asn Leu His Arg Gln Thr Val Leu His Arg Phe Leu Gly Arg  
65 70 75 80Asn Asn Ala Lys Thr Pro Tyr Ile Ile Ser Ile Ala Gly Ser Val Ala  
85 90 95Val Gly Lys Ser Thr Ser Ala Arg Ile Leu Gln Ser Leu Leu Ser His  
100 105 110Trp Pro Thr Glu Arg Lys Val Asp Leu Ile Thr Thr Asp Gly Phe Leu  
115 120 125



- 2 -

Tyr Pro Leu Asn Lys Leu Lys Gln Asp Asn Leu Leu Gln Lys Lys Gly  
 130 135 140  
 Phe Pro Val Ser Tyr Asp Thr Pro Lys Leu Ile Arg Phe Leu Ala Asp  
 145 150 155 160  
 Val Lys Ser Gly Lys Ser Asn Val Thr Ala Pro Ile Tyr Ser His Leu  
 165 170 175  
 Thr Tyr Asp Ile Ile Pro Asp Lys Phe Asp Val Val Asp Lys Pro Asp  
 180 185 190  
 Ile Leu Ile Leu Glu Gly Leu Asn Val Leu Gln Thr Gly Asn Asn Lys  
 195 200 205  
 Thr Asp Gln Thr Phe Val Ser Asp Phe Val Asp Phe Ser Ile Tyr Val  
 210 215 220  
 Asp Ala Glu Glu Lys Leu Leu Lys Glu Trp Tyr Ile Lys Arg Phe Leu  
 225 230 235 240  
 Lys Phe Arg Glu Ser Ala Phe Asn Asp Pro Asn Ser Tyr Phe Lys His  
 245 250 255  
 Tyr Ala Ser Leu Ser Lys Glu Glu Ala Ile Ala Thr Ala Ser Lys Ile  
 260 265 270  
 Trp Asp Glu Ile Asn Gly Leu Asn Leu Asn Gln Asn Ile Leu Pro Thr  
 275 280 285  
 Arg Glu Arg Ala Asn Leu Ile Leu Lys Lys Gly His Asn His Gln Val  
 290 295 300  
 Glu Leu Ile Lys Leu Arg Lys  
 305 310

<210> 2  
 <211> 316  
 <212> PRT  
 <213> Escherichia coli

<400> 2  
 Met Ser Ile Lys Glu Gln Thr Leu Met Thr Pro Tyr Leu Gln Phe Asp  
 1 5 10 15  
 Arg Asn Gln Trp Ala Ala Leu Arg Asp Ser Val Pro Met Thr Leu Ser  
 20 25 30  
 Glu Asp Glu Ile Ala Arg Leu Lys Gly Ile Asn Glu Asp Leu Ser Leu  
 35 40 45  
 Glu Glu Val Ala Glu Ile Tyr Leu Pro Leu Ser Arg Leu Leu Asn Phe  
 50 55 60  
 Tyr Ile Ser Ser Asn Leu Arg Arg Gln Ala Val Leu Glu Gln Phe Leu  
 65 70 75 80

- 3 -

Gly Thr Asn Gly Gln Arg Ile Pro Tyr Ile Ile Ser Ile Ala Gly Ser  
 85 90 95  
 Val Ala Val Gly Lys Ser Thr Thr Ala Arg Val Leu Gln Ala Leu Leu  
 100 105 110  
 Ser Arg Trp Pro Glu His Arg Arg Val Glu Leu Ile Thr Thr Asp Gly  
 115 120 125  
 Phe Leu His Pro Asn Gln Val Leu Lys Glu Arg Gly Leu Met Lys Lys  
 130 135 140  
 Lys Gly Phe Pro Glu Ser Tyr Asp Met His Arg Leu Val Lys Phe Val  
 145 150 155 160  
 Ser Asp Leu Lys Ser Gly Val Pro Asn Val Thr Ala Pro Val Tyr Ser  
 165 170 175  
 His Leu Ile Tyr Asp Val Ile Pro Asp Gly Asp Lys Thr Val Val Gln  
 180 185 190  
 Pro Asp Ile Leu Ile Leu Glu Gly Leu Asn Val Leu Gln Ser Gly Met  
 195 200 205  
 Asp Tyr Pro His Asp Pro His His Val Phe Val Ser Asp Phe Val Asp  
 210 215 220  
 Phe Ser Ile Tyr Val Asp Ala Pro Glu Asp Leu Leu Gln Thr Trp Tyr  
 225 230 235 240  
 Ile Asn Arg Phe Leu Lys Phe Arg Glu Gly Ala Phe Thr Asp Pro Asp  
 245 250 255  
 Ser Tyr Phe His Asn Tyr Ala Lys Leu Thr Lys Glu Glu Ala Ile Lys  
 260 265 270  
 Thr Ala Met Thr Leu Trp Lys Glu Ile Asn Trp Leu Asn Leu Lys Gln  
 275 280 285  
 Asn Ile Leu Pro Thr Arg Glu Arg Ala Ser Leu Ile Leu Thr Lys Ser  
 290 295 300  
 Ala Asn His Ala Val Glu Glu Val Arg Leu Arg Lys  
 305 310 315

&lt;210&gt; 3

&lt;211&gt; 319

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 3

Met Lys Asn Lys Glu Leu Asn Leu His Thr Leu Tyr Thr Gln His Asn  
 1 5 10 15

Arg Glu Ser Trp Ser Gly Phe Gly Gly His Leu Ser Ile Ala Val Ser  
 20 25 30

- 4 -

Glu Glu Glu Ala Lys Ala Val Glu Gly Leu Asn Asp Tyr Leu Ser Val  
 35 40 45  
 Glu Glu Val Glu Thr Ile Tyr Ile Pro Leu Val Arg Leu Leu His Leu  
 50 55 60  
 His Val Lys Ser Ala Ala Glu Arg Asn Lys His Val Asn Val Phe Leu  
 65 70 75 80  
 Lys His Pro His Ser Ala Lys Ile Pro Phe Ile Ile Gly Ile Ala Gly  
 85 90 95  
 Ser Val Ala Val Gly Lys Ser Thr Thr Ala Arg Ile Leu Gln Lys Leu  
 100 105 110  
 Leu Ser Arg Leu Pro Asp Arg Pro Lys Val Ser Leu Ile Thr Thr Asp  
 115 120 125  
 Gly Phe Leu Phe Pro Thr Ala Glu Leu Lys Lys Lys Asn Met Met Ser  
 130 135 140  
 Arg Lys Gly Phe Pro Glu Ser Tyr Asp Val Lys Ala Leu Leu Glu Phe  
 145 150 155 160  
 Leu Asn Asp Leu Lys Ser Gly Lys Asp Ser Val Lys Ala Pro Val Tyr  
 165 170 175  
 Ser His Leu Thr Tyr Asp Arg Glu Glu Gly Val Phe Glu Val Val Glu  
 180 185 190  
 Gln Ala Asp Ile Val Ile Ile Glu Gly Ile Asn Val Leu Gln Ser Pro  
 195 200 205  
 Thr Leu Glu Asp Asp Arg Glu Asn Pro Arg Ile Phe Val Ser Asp Phe  
 210 215 220  
 Phe Asp Phe Ser Ile Tyr Val Asp Ala Glu Glu Ser Arg Ile Phe Thr  
 225 230 235 240  
 Trp Tyr Leu Glu Arg Phe Arg Leu Leu Arg Glu Thr Ala Phe Gln Asn  
 245 250 255  
 Pro Asp Ser Tyr Phe His Lys Phe Lys Asp Leu Ser Asp Gln Glu Ala  
 260 265 270  
 Asp Glu Met Ala Ala Ser Ile Trp Glu Ser Val Asn Arg Pro Asn Leu  
 275 280 285  
 Tyr Glu Asn Ile Leu Pro Thr Lys Phe Arg Ser Asp Leu Ile Leu Arg  
 290 295 300  
 Lys Gly Asp Gly His Lys Val Glu Glu Val Leu Val Arg Arg Val  
 305 310 315

<210> 4  
 <211> 312  
 <212> PRT

- 5 -

&lt;213&gt; Mycobacterium leprae

&lt;400&gt; 4

Met Pro Arg Leu Ser Glu Pro Ser Pro Tyr Val Glu Phe Asp Arg Lys  
 1 5 10 15  
 Gln Trp Arg Ala Leu Arg Met Ser Thr Pro Leu Ala Leu Thr Glu Glu  
 20 25 30  
 Glu Leu Ile Gly Leu Arg Gly Leu Gly Glu Gln Ile Asp Leu Leu Glu  
 35 40 45  
 Val Glu Glu Val Tyr Leu Pro Leu Ala Arg Leu Ile His Leu Gln Val  
 50 55 60  
 Ala Ala Arg Gln Arg Leu Phe Ala Ala Thr Ala Glu Phe Leu Gly Glu  
 65 70 75 80  
 Pro Gln Gln Asn Pro Gly Arg Pro Val Pro Phe Ile Ile Gly Val Ala  
 85 90 95  
 Gly Ser Val Ala Val Gly Lys Ser Thr Thr Ala Arg Val Leu Gln Ala  
 100 105 110  
 Leu Leu Ala Arg Trp Asp His His Thr Arg Val Asp Leu Val Thr Thr  
 115 120 125  
 Asp Gly Phe Leu Tyr Pro Asn Ala Glu Leu Gly Arg Arg Asn Leu Met  
 130 135 140  
 His Arg Lys Gly Phe Pro Glu Ser Tyr Asn Arg Arg Ala Leu Met Arg  
 145 150 155 160  
 Phe Val Thr Ser Val Lys Ser Gly Ala Asp Tyr Ala Cys Ala Pro Val  
 165 170 175  
 Tyr Ser His Leu Arg Tyr Asp Thr Ile Pro Gly Ala Lys His Val Val  
 180 185 190  
 Arg His Pro Asp Ile Leu Ile Leu Glu Gly Leu Asn Val Leu Gln Thr  
 195 200 205  
 Gly Pro Thr Leu Met Val Ser Asp Leu Phe Asp Phe Ser Leu Tyr Val  
 210 215 220  
 Asp Ala Arg Ile Gln Asp Ile Glu Gln Trp Tyr Val Ser Arg Phe Leu  
 225 230 235 240  
 Ala Met Arg Gly Thr Ala Phe Ala Asp Pro Glu Ser His Phe His His  
 245 250 255  
 Tyr Ser Ala Leu Thr Asp Ser Lys Ala Ile Ile Ala Ala Arg Glu Ile  
 260 265 270  
 Trp Arg Ser Ile Asn Arg Pro Asn Leu Val Glu Asn Ile Leu Pro Thr  
 275 280 285  
 Arg Pro Arg Ala Thr Leu Val Leu Arg Lys Asp Ala Asp His Ser Ile

- 6 -

290                      295                      300  
 Asn Arg Leu Arg Leu Arg Lys Leu  
 305                      310  
  
 <210> 5  
 <211> 312  
 <212> PRT  
 <213> Mycobacterium tuberculosis  
  
 <400> 5  
 Met Ser Arg Leu Ser Glu Pro Ser Pro Tyr Val Glu Phe Asp Arg Arg  
 1                      5                      10                      15  
  
 Gln Trp Arg Ala Leu Arg Met Ser Thr Pro Leu Ala Leu Thr Glu Glu  
                     20                      25                      30  
  
 Glu Leu Val Gly Leu Arg Gly Leu Gly Glu Gln Ile Asp Leu Leu Glu  
                     35                      40                      45  
  
 Val Glu Glu Val Tyr Leu Pro Leu Ala Arg Leu Ile His Leu Gln Val  
                     50                      55                      60  
  
 Ala Ala Arg Gln Arg Leu Phe Ala Ala Thr Ala Glu Phe Leu Gly Glu  
                     65                      70                      75                      80  
  
 Pro Gln Gln Asn Pro Asp Arg Pro Val Pro Phe Ile Ile Gly Val Ala  
                     85                      90                      95  
  
 Gly Ser Val Ala Val Gly Lys Ser Thr Thr Ala Arg Val Leu Gln Ala  
                     100                      105                      110  
  
 Leu Leu Ala Arg Trp Asp His His Pro Arg Val Asp Leu Val Thr Thr  
                     115                      120                      125  
  
 Asp Gly Phe Leu Tyr Pro Asn Ala Glu Leu Gln Arg Arg Asn Leu Met  
                     130                      135                      140  
  
 His Arg Lys Gly Phe Pro Glu Ser Tyr Asn Arg Arg Ala Leu Met Arg  
                     145                      150                      155                      160  
  
 Phe Val Thr Ser Val Lys Ser Gly Ser Asp Tyr Ala Cys Ala Pro Val  
                     165                      170                      175  
  
 Tyr Ser His Leu His Tyr Asp Ile Ile Pro Gly Ala Glu Gln Val Val  
                     180                      185                      190  
  
 Arg His Pro Asp Ile Leu Ile Leu Glu Gly Leu Asn Val Leu Gln Thr  
                     195                      200                      205  
  
 Gly Pro Thr Leu Met Val Ser Asp Leu Phe Asp Phe Ser Leu Tyr Val  
                     210                      215                      220  
  
 Asp Ala Arg Ile Glu Asp Ile Glu Gln Trp Tyr Val Ser Arg Phe Leu  
                     225                      230                      235                      240  
  
 Ala Met Arg Thr Thr Ala Phe Ala Asp Pro Glu Ser His Phe His His

- 7 -

				245						250						255
Tyr	Ala	Ala	Phe	Ser	Asp	Ser	Gln	Ala	Val	Val	Ala	Ala	Arg	Glu	Ile	
			260					265					270			
Trp	Arg	Thr	Ile	Asn	Arg	Pro	Asn	Leu	Val	Glu	Asn	Ile	Leu	Pro	Thr	
		275					280					285				
Arg	Pro	Arg	Ala	Thr	Leu	Val	Leu	Arg	Lys	Asp	Ala	Asp	His	Ser	Ile	
	290					295					300					
Asn	Arg	Leu	Arg	Leu	Arg	Lys	Leu									
305					310											

&lt;210&gt; 6

&lt;211&gt; 329

&lt;212&gt; PRT

&lt;213&gt; Streptomyces coelicolor

&lt;400&gt; 6

Met	Ile	Ser	Pro	Val	Pro	Ser	Ile	Pro	Arg	Ser	Ala	His	Arg	Gln	Arg
1				5					10					15	
Pro	Glu	Ala	Thr	Pro	Tyr	Val	Asp	Leu	Thr	Arg	Pro	Glu	Trp	Ser	Ala
			20					25					30		
Leu	Arg	Asp	Lys	Thr	Pro	Leu	Pro	Leu	Thr	Ala	Glu	Glu	Val	Glu	Lys
		35					40					45			
Leu	Arg	Gly	Leu	Gly	Asp	Val	Ile	Asp	Leu	Asp	Glu	Val	Arg	Asp	Ile
	50					55					60				
Tyr	Leu	Pro	Leu	Ser	Arg	Leu	Leu	Asn	Leu	Tyr	Val	Gly	Ala	Thr	Asp
65					70					75					80
Gly	Leu	Arg	Gly	Ala	Leu	Asn	Thr	Phe	Leu	Gly	Glu	Gln	Gly	Ser	Gln
				85					90					95	
Ser	Gly	Thr	Pro	Phe	Val	Ile	Gly	Val	Ala	Gly	Ser	Val	Ala	Val	Gly
			100					105					110		
Lys	Ser	Thr	Val	Ala	Arg	Leu	Leu	Gln	Ala	Leu	Leu	Ser	Arg	Trp	Pro
		115					120					125			
Glu	His	Pro	Arg	Val	Glu	Leu	Val	Thr	Thr	Asp	Gly	Phe	Leu	Leu	Pro
	130					135					140				
Thr	Arg	Glu	Leu	Glu	Ala	Arg	Gly	Leu	Met	Ser	Arg	Lys	Gly	Phe	Pro
145					150					155					160
Glu	Ser	Tyr	Asp	Arg	Arg	Ala	Leu	Thr	Arg	Phe	Val	Ala	Asp	Ile	Lys
				165					170					175	
Ala	Gly	Lys	Ala	Glu	Val	Thr	Ala	Pro	Val	Tyr	Ser	His	Leu	Ile	Tyr
			180					185					190		
Asp	Ile	Val	Pro	Asp	Gln	Arg	Leu	Val	Val	Arg	Arg	Pro	Asp	Ile	Leu

- 8 -

195					200					205					
Ile	Val	Glu	Gly	Leu	Asn	Val	Leu	Gln	Pro	Ala	Leu	Pro	Gly	Lys	Asp
210					215					220					
Gly	Arg	Thr	Arg	Val	Gly	Leu	Ala	Asp	Tyr	Phe	Asp	Phe	Ser	Val	Tyr
225					230					235					240
Val	Asp	Ala	Arg	Thr	Glu	Asp	Ile	Glu	Arg	Trp	Tyr	Leu	Asn	Arg	Phe
				245					250					255	
Arg	Lys	Leu	Arg	Ala	Thr	Ala	Phe	Gln	Asn	Pro	Ser	Ser	Tyr	Phe	Arg
			260					265					270		
Lys	Tyr	Thr	Gln	Val	Ser	Glu	Glu	Glu	Ala	Leu	Asp	Tyr	Ala	Arg	Thr
		275					280					285			
Thr	Trp	Arg	Thr	Ile	Asn	Lys	Pro	Asn	Leu	Val	Glu	Asn	Val	Ala	Pro
	290					295					300				
Thr	Arg	Gly	Arg	Ala	Thr	Leu	Val	Leu	Arg	Lys	Gly	Pro	Asp	His	Lys
305					310					315					320
Val	Gln	Arg	Leu	Ser	Leu	Arg	Lys	Leu							
				325											

&lt;210&gt; 7

&lt;211&gt; 265

&lt;212&gt; PRT

&lt;213&gt; Streptomyces coelicolor

&lt;400&gt; 7

Met	Leu	Leu	Thr	Ile	Asp	Val	Gly	Asn	Thr	His	Thr	Val	Leu	Gly	Leu
1				5					10					15	
Phe	Asp	Gly	Glu	Asp	Ile	Val	Glu	His	Trp	Arg	Ile	Ser	Thr	Asp	Ser
			20					25					30		
Arg	Arg	Thr	Ala	Asp	Glu	Leu	Ala	Val	Leu	Leu	Gln	Gly	Leu	Met	Gly
			35				40					45			
Met	His	Pro	Leu	Leu	Gly	Asp	Glu	Leu	Gly	Asp	Gly	Ile	Asp	Gly	Ile
	50					55					60				
Ala	Ile	Cys	Ala	Thr	Val	Pro	Ser	Val	Leu	His	Glu	Leu	Arg	Glu	Val
	65				70						75				80
Thr	Arg	Arg	Tyr	Tyr	Gly	Asp	Val	Pro	Ala	Val	Leu	Val	Glu	Pro	Gly
			85					90					95		
Val	Lys	Thr	Gly	Val	Pro	Ile	Leu	Thr	Asp	His	Pro	Lys	Glu	Val	Gly
			100					105					110		
Ala	Asp	Arg	Ile	Ile	Asn	Ala	Val	Ala	Ala	Val	Glu	Leu	Tyr	Gly	Gly
		115				120						125			
Pro	Ala	Ile	Val	Val	Asp	Phe	Gly	Thr	Ala	Thr	Thr	Phe	Asp	Ala	Val

- 9 -

130						135						140			
Ser	Ala	Arg	Gly	Glu	Tyr	Ile	Gly	Gly	Val	Ile	Ala	Pro	Gly	Ile	Glu
145					150					155					160
Ile	Ser	Val	Glu	Ala	Leu	Gly	Val	Lys	Gly	Ala	Gln	Leu	Arg	Lys	Ile
				165					170					175	
Glu	Val	Ala	Arg	Pro	Arg	Ser	Val	Ile	Gly	Lys	Asn	Thr	Val	Glu	Ala
			180					185					190		
Met	Gln	Ser	Gly	Ile	Val	Tyr	Gly	Phe	Ala	Gly	Gln	Val	Asp	Gly	Val
		195					200					205			
Val	Asn	Arg	Met	Ala	Arg	Glu	Leu	Ala	Asp	Asp	Pro	Asp	Asp	Val	Thr
	210					215					220				
Val	Ile	Ala	Thr	Gly	Gly	Leu	Ala	Pro	Met	Val	Leu	Gly	Glu	Ser	Ser
225					230					235					240
Val	Ile	Asp	Glu	His	Glu	Pro	Trp	Leu	Thr	Leu	Met	Gly	Leu	Arg	Leu
				245					250					255	
Val	Tyr	Glu	Arg	Asn	Val	Ser	Arg	Met							
			260					265							

&lt;210&gt; 8

&lt;211&gt; 272

&lt;212&gt; PRT

&lt;213&gt; Mycobacterium tuberculosis

&lt;400&gt; 8

Met	Leu	Leu	Ala	Ile	Asp	Val	Arg	Asn	Thr	His	Thr	Val	Val	Gly	Leu
1				5					10					15	
Leu	Ser	Gly	Met	Lys	Glu	His	Ala	Lys	Val	Val	Gln	Gln	Trp	Arg	Ile
			20					25					30		
Arg	Thr	Glu	Ser	Glu	Val	Thr	Ala	Asp	Glu	Leu	Ala	Leu	Thr	Ile	Asp
		35					40					45			
Gly	Leu	Ile	Gly	Glu	Asp	Ser	Glu	Arg	Leu	Thr	Gly	Thr	Ala	Ala	Leu
	50					55					60				
Ser	Thr	Val	Pro	Ser	Val	Leu	His	Glu	Val	Arg	Ile	Met	Leu	Asp	Gln
	65				70					75				80	
Tyr	Trp	Pro	Ser	Val	Pro	His	Val	Leu	Ile	Glu	Pro	Gly	Val	Arg	Thr
				85					90					95	
Gly	Ile	Pro	Leu	Leu	Val	Asp	Asn	Pro	Lys	Glu	Val	Gly	Ala	Asp	Arg
			100					105					110		
Ile	Val	Asn	Cys	Leu	Ala	Ala	Tyr	Asp	Arg	Phe	Arg	Lys	Ala	Ala	Ile
		115					120					125			
Val	Val	Asp	Phe	Gly	Ser	Ser	Ile	Cys	Val	Asp	Val	Val	Ser	Ala	Lys



- 10 -

130                      135                      140  
 Gly Glu Phe Leu Gly Gly Ala Ile Ala Pro Gly Val Gln Val Ser Ser  
 145                      150                      155                      160  
 Asp Ala Ala Ala Ala Arg Ser Ala Ala Leu Arg Arg Val Glu Leu Ala  
                     165                      170                      175  
 Arg Pro Arg Ser Val Val Gly Lys Asn Thr Val Glu Cys Met Gln Ala  
                     180                      185                      190  
 Gly Ala Val Phe Gly Phe Ala Gly Leu Val Asp Gly Leu Val Gly Arg  
                     195                      200                      205  
 Ile Arg Glu Asp Val Ser Gly Phe Ser Val Asp His Asp Val Ala Ile  
                     210                      215                      220  
 Val Ala Thr Gly His Thr Ala Pro Leu Leu Leu Pro Glu Leu His Thr  
 225                      230                      235                      240  
 Val Asp His Tyr Asp Gln His Leu Thr Leu Gln Gly Leu Arg Leu Val  
                     245                      250                      255  
 Phe Glu Arg Asn Leu Glu Val Gln Arg Gly Arg Leu Lys Thr Ala Arg  
                     260                      265                      270

&lt;210&gt; 9

&lt;211&gt; 258

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 9

Leu Leu Leu Val Ile Asp Val Gly Asn Thr Asn Thr Val Leu Gly Val  
 1                      5                      10                      15  
 Tyr His Asp Gly Lys Leu Glu Tyr His Trp Arg Ile Glu Thr Ser Arg  
                     20                      25                      30  
 His Lys Thr Glu Asp Glu Phe Gly Met Ile Leu Arg Ser Leu Phe Asp  
                     35                      40                      45  
 His Ser Gly Leu Met Phe Glu Gln Ile Asp Gly Ile Ile Ile Ser Ser  
                     50                      55                      60  
 Val Val Pro Pro Ile Met Phe Ala Leu Glu Arg Met Cys Thr Lys Tyr  
 65                      70                      75                      80  
 Phe His Ile Glu Pro Gln Ile Val Gly Pro Gly Met Lys Thr Gly Leu  
                     85                      90                      95  
 Asn Ile Lys Tyr Asp Asn Pro Lys Glu Val Gly Ala Asp Arg Ile Val  
                     100                      105                      110  
 Asn Ala Val Ala Ala Ile His Leu Tyr Gly Asn Pro Leu Ile Val Val  
                     115                      120                      125

- 11 -

Asp Phe Gly Thr Ala Thr Thr Tyr Cys Tyr Ile Asp Glu Asn Lys Gln  
 130 135 140  
 Tyr Met Gly Gly Ala Ile Ala Pro Gly Ile Thr Ile Ser Thr Glu Ala  
 145 150 155 160  
 Leu Tyr Ser Arg Ala Ala Lys Leu Pro Arg Ile Glu Ile Thr Arg Pro  
 165 170 175  
 Asp Asn Ile Ile Gly Lys Asn Thr Val Ser Ala Met Gln Ser Gly Ile  
 180 185 190  
 Leu Phe Gly Tyr Val Gly Gln Val Glu Gly Ile Val Lys Arg Met Lys  
 195 200 205  
 Trp Gln Ala Lys Gln Asp Leu Lys Val Ile Ala Thr Gly Gly Leu Ala  
 210 215 220  
 Pro Leu Ile Ala Asn Glu Ser Asp Cys Ile Asp Ile Val Asp Pro Phe  
 225 230 235 240  
 Leu Thr Leu Lys Gly Leu Glu Leu Ile Tyr Glu Arg Asn Arg Val Gly  
 245 250 255  
 Ser Val

<210> 10  
 <211> 262  
 <212> PRT  
 <213> Deinococcus radiopugnans

<400> 10  
 Met Pro Ala Phe Pro Leu Leu Ala Val Asp Ile Gly Asn Thr Thr Thr  
 1 5 10 15  
 Val Leu Gly Leu Ala Asp Ala Ser Gly Ala Leu Thr His Thr Trp Arg  
 20 25 30  
 Ile Arg Thr Asn Arg Glu Met Leu Pro Asp Asp Leu Ala Leu Gln Leu  
 35 40 45  
 His Gly Leu Phe Thr Leu Ala Gly Ala Pro Ile Pro Arg Ala Ala Val  
 50 55 60  
 Leu Ser Ser Val Ala Pro Pro Val Gly Glu Asn Tyr Ala Leu Ala Leu  
 65 70 75 80  
 Lys Arg His Phe Met Ile Asp Ala Phe Ala Val Ser Ala Glu Asn Leu  
 85 90 95  
 Pro Asp Val Thr Val Glu Leu Asp Thr Pro Gly Ser Val Gly Ala Asp  
 100 105 110  
 Arg Leu Cys Asn Leu Phe Gly Ala Glu Lys Tyr Leu Gly Gly Leu Asp

- 12 -

115                      120                      125  
 Tyr Ala Val Val Val Asp Phe Gly Thr Ser Thr Asn Phe Asp Val Val  
 130                      135                      140  
 Gly Arg Gly Arg Arg Phe Leu Gly Gly Ile Leu Ala Thr Gly Ala Gln  
 145                      150                      155                      160  
 Val Ser Ala Asp Ala Leu Phe Ala Arg Ala Ala Lys Leu Pro Arg Ile  
 165                      170                      175  
 Thr Leu Gln Ala Pro Glu Thr Ala Ile Gly Lys Asn Thr Val His Ala  
 180                      185                      190  
 Leu Gln Ser Gly Leu Val Phe Gly Tyr Ala Glu Met Val Asp Gly Leu  
 195                      200                      205  
 Leu Arg Arg Ile Arg Ala Glu Leu Pro Gly Glu Ala Val Ala Val Ala  
 210                      215                      220  
 Thr Gly Gly Phe Ser Arg Thr Val Gln Gly Ile Cys Gln Glu Ile Asp  
 225                      230                      235                      240  
 Tyr Tyr Asp Glu Thr Leu Thr Leu Arg Gly Leu Val Glu Leu Trp Ala  
 245                      250                      255  
 Ser Arg Ser Glu Val Arg  
 260

<210> 11  
 <211> 212  
 <212> PRT  
 <213> Desulfovibrio vulgaris

<400> 11  
 Met Thr Gln His Phe Leu Leu Phe Asp Ile Gly Asn Thr Asn Val Lys  
 1                      5                      10                      15  
 Ile Gly Ile Ala Val Glu Thr Ala Val Leu Thr Ser Tyr Val Leu Pro  
 20                      25                      30  
 Thr Asp Pro Gly Gln Thr Thr Asp Ser Ile Gly Leu Arg Leu Leu Glu  
 35                      40                      45  
 Val Leu Arg His Ala Gly Leu Gly Pro Ala Asp Val Gly Ala Cys Val  
 50                      55                      60  
 Ala Ser Ser Val Val Pro Gly Val Asn Pro Leu Ile Arg Arg Ala Cys  
 65                      70                      75                      80  
 Glu Arg Tyr Leu Tyr Arg Lys Leu Leu Phe Ala Pro Gly Asp Ile Ala  
 85                      90                      95  
 Ile Pro Leu Asp Asn Arg Tyr Glu Arg Pro Ala Glu Val Gly Ala Asp  
 100                      105                      110  
 Arg Leu Val Ala Ala Tyr Ala Ala Arg Arg Leu Tyr Pro Gly Pro Arg

- 13 -

115                      120                      125  
 Ser Leu Val Ser Val Asp Phe Gly Thr Ala Thr Thr Phe Asp Cys Val  
 130                      135                      140  
 Glu Gly Gly Ala Tyr Leu Gly Gly Leu Ile Cys Pro Gly Val Leu Ser  
 145                      150                      155                      160  
 Ser Ala Gly Ala Leu Ser Ser Arg Thr Ala Lys Leu Pro Arg Ile Ser  
 165                      170  
 Leu Glu Val Glu Glu Asp Ser Pro Val Ile Gly Arg Ser Thr Thr Thr  
 180                      185                      190  
 Ser Leu Asn His Gly Phe Ile Phe Gly Phe Ala Ala Met Thr Glu Gly  
 195                      200                      205  
 Val Leu Ala Ala  
 210

&lt;210&gt; 12.

&lt;211&gt; 246

&lt;212&gt; PRT

&lt;213&gt; Thermotoga maritima

&lt;400&gt; 12

Met Tyr Leu Leu Val Asp Val Gly Asn Thr His Ser Val Phe Ser Ile  
 1                      5                      10                      15  
 Thr Glu Asp Gly Lys Thr Phe Arg Arg Trp Arg Leu Ser Thr Gly Val  
 20                      25                      30  
 Phe Gln Thr Glu Asp Glu Leu Phe Ser His Leu His Pro Leu Leu Gly  
 35                      40                      45  
 Asp Ala Met Arg Glu Ile Lys Gly Ile Gly Val Ala Ser Val Val Pro  
 50                      55                      60  
 Thr Gln Asn Thr Val Ile Glu Arg Phe Ser Gln Lys Tyr Phe His Ile  
 65                      70                      75                      80  
 Ser Pro Ile Trp Val Lys Ala Lys Asn Gly Cys Val Lys Trp Asn Val  
 85                      90                      95  
 Lys Asn Pro Ser Glu Val Gly Ala Asp Arg Val Ala Asn Val Val Ala  
 100                      105                      110  
 Phe Val Lys Glu Tyr Gly Lys Asn Gly Ile Ile Ile Asp Met Gly Thr  
 115                      120                      125  
 Ala Thr Thr Val Asp Leu Val Val Asn Gly Ser Tyr Glu Gly Gly Ala  
 130                      135                      140  
 Ile Leu Pro Gly Phe Phe Met Met Val His Ser Leu Phe Arg Gly Thr  
 145                      150                      155                      160  
 Ala Lys Leu Pro Leu Val Glu Val Lys Pro Ala Asp Phe Val Val Gly

- 14 -

165 170 175  
 Lys Asp Thr Glu Glu Asn Ile Arg Leu Gly Val Val Asn Gly Ser Val  
 180 185 190  
 Tyr Ala Leu Glu Gly Ile Ile Gly Arg Ile Lys Glu Val Tyr Gly Asp  
 195 200 205  
 Leu Pro Val Val Leu Thr Gly Gly Gln Ser Lys Ile Val Lys Asp Met  
 210 215 220  
 Ile Lys His Glu Ile Phe Asp Glu Asp Leu Thr Ile Lys Gly Val Tyr  
 225 230 235 240  
 His Phe Cys Phe Gly Asp  
 245

&lt;210&gt; 13

&lt;211&gt; 273

&lt;212&gt; PRT

<213> *Treponema pallidum*

&lt;400&gt; 13

Met Leu Leu Ile Asp Val Gly Asn Ser His Val Val Phe Gly Ile Gln  
 1 5 10 15

Gly Glu Asn Gly Gly Arg Val Cys Val Arg Glu Leu Phe Arg Leu Ala  
 20 25 30

Pro Asp Ala Arg Lys Thr Gln Asp Glu Tyr Ser Leu Leu Ile His Ala  
 35 40 45

Leu Cys Glu Arg Ala Gly Val Gly Arg Ala Ser Leu Arg Asp Ala Phe  
 50 55 60

Ile Ser Ser Val Val Pro Val Leu Thr Lys Thr Ile Ala Asp Ala Val  
 65 70 75 80

Ala Gln Ile Ser Gly Val Gln Pro Val Val Phe Gly Pro Trp Ala Tyr  
 85 90 95

Glu His Leu Pro Val Arg Ile Pro Glu Pro Val Arg Ala Glu Ile Gly  
 100 105 110

Thr Asp Leu Val Ala Asn Ala Val Ala Ala Tyr Val His Phe Arg Ser  
 115 120 125

Ala Cys Val Val Val Asp Cys Gly Thr Ala Leu Thr Phe Thr Ala Val  
 130 135 140

Asp Gly Thr Gly Leu Ile Gln Gly Val Ala Ile Ala Pro Gly Leu Arg  
 145 150 155 160

Thr Ala Val Gln Ser Leu His Thr Gly Thr Ala Gln Leu Pro Leu Val  
 165 170 175

Pro Leu Ala Leu Pro Asp Ser Val Leu Gly Lys Asp Thr Thr His Ala

- 15 -

			180						185					190			
Val	Gln	Ala	Gly	Val	Val	Arg	Gly	Thr	Leu	Phe	Val	Ile	Arg	Ala	Met		
		195					200					205					
Ile	Ala	Gln	Cys	Gln	Lys	Glu	Leu	Gly	Cys	Arg	Cys	Ala	Ala	Val	Ile		
	210					215					220						
Thr	Gly	Gly	Leu	Ser	Arg	Leu	Phe	Ser	Ser	Glu	Val	Asp	Phe	Pro	Pro		
225					230					235					240		
Ile	Asp	Ala	Gln	Leu	Thr	Leu	Ser	Gly	Leu	Ala	His	Ile	Ala	Arg	Leu		
				245					250					255			
Val	Pro	Thr	Ser	Leu	Leu	Pro	Pro	Ala	Thr	Val	Ser	Gly	Ser	Ser	Gly		
			260					265					270				

Asn

&lt;210&gt; 14

&lt;211&gt; 262

&lt;212&gt; PRT

&lt;213&gt; Borrelia burgdorferi

&lt;400&gt; 14

Met	Asn	Lys	Pro	Leu	Leu	Ser	Glu	Leu	Ile	Ile	Asp	Ile	Gly	Asn	Thr
1				5					10					15	
Ser	Ile	Ala	Phe	Ala	Leu	Phe	Lys	Asp	Asn	Gln	Val	Asn	Leu	Phe	Ile
		20					25						30		
Lys	Met	Lys	Thr	Asn	Leu	Met	Leu	Arg	Tyr	Asp	Glu	Val	Tyr	Ser	Phe
		35					40					45			
Phe	Glu	Glu	Asn	Phe	Asp	Phe	Asn	Val	Asn	Lys	Val	Phe	Ile	Ser	Ser
	50					55					60				
Val	Val	Pro	Ile	Leu	Asn	Glu	Thr	Phe	Lys	Asn	Val	Ile	Phe	Ser	Phe
65					70					75					80
Phe	Lys	Ile	Lys	Pro	Leu	Phe	Ile	Gly	Phe	Asp	Leu	Asn	Tyr	Asp	Leu
				85					90					95	
Thr	Phe	Asn	Pro	Tyr	Lys	Ser	Asp	Lys	Phe	Leu	Leu	Gly	Ser	Asp	Val
			100					105					110		
Phe	Ala	Asn	Leu	Val	Ala	Ala	Ile	Glu	Asn	Tyr	Ser	Phe	Glu	Asn	Val
		115					120					125			
Leu	Val	Val	Asp	Leu	Gly	Thr	Ala	Cys	Thr	Ile	Phe	Ala	Val	Ser	Arg
	130					135					140				
Gln	Asp	Gly	Ile	Leu	Gly	Gly	Ile	Ile	Asn	Ser	Gly	Pro	Leu	Ile	Asn
145					150					155					160
Phe	Asn	Ser	Leu	Leu	Asp	Asn	Ala	Tyr	Leu	Ile	Lys	Lys	Phe	Pro	Ile

- 16 -

165                      170                      175  
 Ser Thr Pro Asn Asn Leu Leu Glu Arg Thr Thr Ser Gly Ser Val Asn  
                          180                      185                      190  
 Ser Gly Leu Phe Tyr Gln Tyr Lys Tyr Leu Ile Glu Gly Val Tyr Arg  
                          195                      200                      205  
 Asp Ile Lys Gln Met Tyr Lys Lys Lys Phe Asn Leu Ile Ile Thr Gly  
                          210                      215                      220  
 Gly Asn Ala Asp Leu Ile Leu Ser Leu Ile Glu Ile Glu Phe Ile Phe  
                          225                      230                      235                      240  
 Asn Ile His Leu Thr Val Glu Gly Val Arg Ile Leu Gly Asn Ser Ile  
                          245                      250                      255  
 Asp Phe Lys Phe Val Asn  
                          260

<210> 15  
 <211> 229  
 <212> PRT  
 <213> Aquifex aeolicus

<400> 15  
 Met Arg Phe Leu Thr Val Asp Val Gly Asn Ser Ser Val Asp Ile Ala  
                          1                      5                      10                      15  
 Leu Trp Glu Gly Lys Lys Val Lys Asp Phe Leu Lys Leu Ser His Glu  
                          20                      25                      30  
 Glu Phe Leu Lys Glu Glu Phe Pro Lys Leu Lys Ala Leu Gly Ile Ser  
                          35                      40                      45  
 Val Lys Gln Ser Phe Ser Glu Lys Val Arg Gly Lys Ile Pro Lys Ile  
                          50                      55                      60  
 Lys Phe Leu Lys Lys Glu Asn Phe Pro Ile Gln Val Asp Tyr Lys Thr  
                          65                      70                      75                      80  
 Pro Glu Thr Leu Gly Thr Asp Arg Val Ala Leu Ala Tyr Ser Ala Lys  
                          85                      90                      95  
 Lys Phe Tyr Gly Lys Asn Val Val Val Ile Ser Ala Gly Thr Ala Leu  
                          100                      105                      110  
 Val Ile Asp Leu Val Leu Glu Gly Lys Phe Lys Gly Gly Phe Ile Thr  
                          115                      120                      125  
 Leu Gly Leu Gly Lys Lys Leu Lys Ile Leu Ser Asp Leu Ala Glu Gly  
                          130                      135                      140  
 Ile Pro Glu Phe Phe Pro Glu Glu Val Glu Ile Phe Leu Gly Arg Ser  
                          145                      150                      155                      160  
 Thr Arg Glu Cys Val Leu Gly Gly Ala Tyr Arg Glu Ser Thr Glu Phe

- 17 -

165 170 175  
 Ile Lys Ser Thr Leu Lys Leu Trp Arg Lys Val Phe Lys Arg Lys Phe  
 180 185 190  
 Lys Val Val Ile Thr Gly Gly Glu Gly Lys Tyr Phe Ser Lys Phe Gly  
 195 200 205  
 Ile Tyr Asp Pro Leu Leu Val His Arg Gly Met Arg Asn Leu Leu Tyr  
 210 215 220  
 Leu Tyr His Arg Ile  
 225

<210> 16  
 <211> 257  
 <212> PRT  
 <213> Synechocystis sp.

<400> 16  
 Met Glu Thr Ser Lys Pro Gly Cys Gly Leu Ala Leu Asp Asn Asp Lys  
 1 5 10 15  
 Gln Lys Pro Trp Leu Gly Leu Met Ile Gly Asn Ser Arg Leu His Trp  
 20 25 30  
 Ala Tyr Cys Ser Gly Asn Ala Pro Leu Gln Thr Trp Val Thr Asp Tyr  
 35 40 45  
 Asn Pro Lys Ser Ala Gln Leu Pro Val Leu Leu Gly Lys Val Pro Leu  
 50 55 60  
 Met Leu Ala Ser Val Val Pro Glu Gln Thr Glu Val Trp Arg Val Tyr  
 65 70 75 80  
 Gln Pro Lys Ile Leu Thr Leu Lys Asn Leu Pro Leu Val Asn Leu Tyr  
 85 90 95  
 Pro Ser Phe Gly Ile Asp Arg Ala Leu Ala Gly Leu Gly Thr Gly Leu  
 100 105 110  
 Thr Tyr Gly Phe Pro Cys Leu Val Val Asp Gly Gly Thr Ala Leu Thr  
 115 120 125  
 Ile Thr Gly Phe Asp Gln Asp Lys Lys Leu Val Gly Gly Ala Ile Leu  
 130 135 140  
 Pro Gly Leu Gly Leu Gln Leu Ala Thr Leu Gly Asp Arg Leu Ala Ala  
 145 150 155 160  
 Leu Pro Lys Leu Glu Met Asp Gln Leu Thr Glu Leu Pro Asp Arg Trp  
 165 170 175  
 Ala Leu Asp Thr Pro Ser Ala Ile Phe Ser Gly Val Val Tyr Gly Val  
 180 185 190  
 Leu Gly Ala Leu Gln Ser Tyr Leu Gln Asp Trp Gln Lys Leu Phe Pro



- 18 -

195	200	205
Gly Ala Ala Met Val Ile Thr	Gly Gly Asp Gly Lys	Ile Leu His Gly
210	215	220
Phe Leu Lys Glu His Ser Pro	Asn Leu Ser Val Ala Trp Asp Asp Asn	
225	230	235 240
Leu Ile Phe Leu Gly Met Ala	Ala Ile His His Gly Asp Arg Pro Ile	
	245	250 255

Cys

<210> 17  
 <211> 223  
 <212> PRT  
 <213> Helicobacter pylori

&lt;400&gt; 17

Met Pro Ala Arg Gln Ser Phe	Thr Asp Leu Lys Asn Leu Val Leu Cys
1	5 10 15
Asp Ile Gly Asn Thr Arg Ile His Phe Ala Gln Asn Tyr Gln Leu Phe	
	20 25 30
Ser Ser Ala Lys Glu Asp Leu Lys Arg Leu Gly Ile Gln Lys Glu Ile	
	35 40 45
Phe Tyr Ile Ser Val Asn Glu Glu Asn Glu Lys Ala Leu Leu Asn Cys	
	50 55 60
Tyr Pro Asn Ala Lys Asn Ile Ala Gly Phe Phe His Leu Glu Thr Asp	
	65 70 75 80
Tyr Val Gly Leu Gly Ile Asp Arg Gln Met Ala Cys Leu Ala Val Asn	
	85 90 95
Asn Gly Val Val Val Asp Ala Gly Ser Ala Ile Thr Ile Asp Leu Ile	
	100 105 110
Lys Glu Gly Lys His Leu Gly Gly Cys Ile Leu Pro Gly Leu Ala Gln	
	115 120 125
Tyr Ile His Ala Tyr Lys Lys Ser Ala Lys Ile Leu Glu Gln Pro Phe	
	130 135 140
Lys Ala Leu Asp Ser Leu Glu Val Leu Pro Lys Ser Thr Arg Asp Ala	
	145 150 155 160
Val Asn Tyr Gly Met Val Leu Ser Val Ile Ala Cys Ile Gln His Leu	
	165 170 175
Ala Lys Asn Gln Lys Ile Tyr Leu Cys Gly Gly Asp Ala Lys Tyr Leu	
	180 185 190
Ser Ala Phe Leu Pro His Ser Val Cys Lys Glu Arg Leu Val Phe Asp	

- 19 -

195                      200                      205  
 Gly Met Glu Ile Ala Leu Lys Lys Ala Gly Ile Leu Glu Cys Lys  
       210                      215                      220

<210> 18  
 <211> 267  
 <212> PRT  
 <213> Bordetella pertussis

<400> 18  
 Met Ile Ile Leu Ile Asp Ser Gly Asn Ser Arg Leu Lys Val Gly Trp  
       1                      5                      10                      15

Phe Asp Pro Asp Ala Pro Gln Ala Ala Arg Glu Pro Ala Pro Val Ala  
                     20                      25                      30

Phe Asp Asn Leu Asp Leu Asp Ala Leu Gly Arg Trp Leu Ala Thr Leu  
                     35                      40                      45

Pro Arg Arg Pro Gln Arg Ala Leu Gly Val Asn Val Ala Gly Leu Ala  
                     50                      55                      60

Arg Gly Glu Ala Ile Ala Ala Thr Leu Arg Ala Gly Gly Cys Asp Ile  
       65                      70                      75                      80

Arg Trp Leu Arg Ala Gln Pro Leu Ala Met Gly Leu Arg Asn Gly Tyr  
                     85                      90                      95

Arg Asn Pro Asp Gln Leu Gly Ala Asp Arg Trp Ala Cys Met Val Gly  
                     100                      105                      110

Val Leu Ala Arg Gln Pro Ser Val His Pro Pro Leu Leu Val Ala Ser  
                     115                      120                      125

Phe Gly Thr Ala Thr Thr Leu Asp Thr Ile Gly Pro Asp Asn Val Phe  
       130                      135                      140

Pro Gly Gly Leu Ile Leu Pro Gly Pro Ala Met Met Arg Gly Ala Leu  
       145                      150                      155                      160

Ala Tyr Gly Thr Ala His Leu Pro Leu Ala Asp Gly Leu Val Ala Asp  
                     165                      170                      175

Tyr Pro Ile Asp Thr His Gln Ala Ile Ala Ser Gly Ile Ala Ala Ala  
                     180                      185                      190

Gln Ala Gly Ala Ile Val Arg Gln Trp Leu Ala Gly Arg Gln Arg Tyr  
                     195                      200                      205

Gly Gln Ala Pro Glu Ile Tyr Val Ala Gly Gly Gly Trp Pro Glu Val  
       210                      215                      220

Arg Gln Glu Ala Glu Arg Leu Leu Ala Val Thr Gly Ala Ala Phe Gly  
       225                      230                      235                      240

Ala Thr Pro Gln Pro Thr Tyr Leu Asp Ser Pro Val Leu Asp Gly Leu

- 20 -

[illegible]

- 21 -

165										170					175					
gac	aat	att	atc	gga	aaa	aac	act	gtt	agc	gcg	atg	caa	tct	gga	att	576				
Asp	Asn	Ile	Ile	Gly	Lys	Asn	Thr	Val	Ser	Ala	Met	Gln	Ser	Gly	Ile					
			180					185					190							
tta	ttt	ggc	tat	gtc	ggc	caa	gtg	gaa	gga	atc	gtt	aag	cga	atg	aaa	624				
Leu	Phe	Gly	Tyr	Val	Gly	Gln	Val	Glu	Gly	Ile	Val	Lys	Arg	Met	Lys					
		195					200					205								
tgg	cag	gca	aaa	cag	gac	ctc	aag	gtc	att	gcg	aca	gga	ggc	ctg	gcg	672				
Trp	Gln	Ala	Lys	Gln	Asp	Leu	Lys	Val	Ile	Ala	Thr	Gly	Gly	Leu	Ala					
	210					215					220									
ccg	ctc	att	gcg	aac	gaa	tca	gat	tgt	ata	gac	atc	gtt	gat	cca	ttc	720				
Pro	Leu	Ile	Ala	Asn	Glu	Ser	Asp	Cys	Ile	Asp	Ile	Val	Asp	Pro	Phe					
	225				230					235					240					
tta	acc	cta	aaa	ggg	ctg	gaa	ttg	att	tat	gaa	aga	aac	cgc	gta	gga	768				
Leu	Thr	Leu	Lys	Gly	Leu	Glu	Leu	Ile	Tyr	Glu	Arg	Asn	Arg	Val	Gly					
			245					250						255						
agt	gta	tag														777				
Ser	Val																			

&lt;210&gt; 20

&lt;211&gt; 960

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(957)

&lt;400&gt; 20

gtg	aaa	aat	aaa	gaa	ctt	aac	cta	cat	act	tta	tat	aca	cag	cac	aat	48
Met	Lys	Asn	Lys	Glu	Leu	Asn	Leu	His	Thr	Leu	Tyr	Thr	Gln	His	Asn	
1				5					10					15		

cgg	gag	tct	tgg	tct	ggt	ttt	ggg	ggg	cat	ttg	tcg	att	gct	gta	tct	96
Arg	Glu	Ser	Trp	Ser	Gly	Phe	Gly	Gly	His	Leu	Ser	Ile	Ala	Val	Ser	
		20					25						30			

gaa	gaa	gag	gca	aaa	gct	gtg	gaa	gga	ttg	aat	gat	tat	cta	tct	gtt	144
Glu	Glu	Glu	Ala	Lys	Ala	Val	Glu	Gly	Leu	Asn	Asp	Tyr	Leu	Ser	Val	
		35				40						45				

gaa	gaa	gtg	gag	acg	atc	tat	att	ccg	ctt	gtt	cgc	ttg	ctt	cat	tta	192
Glu	Glu	Val	Glu	Thr	Ile	Tyr	Ile	Pro	Leu	Val	Arg	Leu	Leu	His	Leu	
		50			55					60						

cat	gtc	aag	tct	gcg	gct	gaa	cgc	aat	aag	cat	gtc	aat	gtt	ttt	ttg	240
His	Val	Lys	Ser	Ala	Ala	Glu	Arg	Asn	Lys	His	Val	Asn	Val	Phe	Leu	
	65				70				75						80	

aag	cac	cca	cat	tca	gcc	aaa	att	ccg	ttt	att	atc	ggc	att	gcc	ggc	288
Lys	His	Pro	His	Ser	Ala	Lys	Ile	Pro	Phe	Ile	Ile	Gly	Ile	Ala	Gly	

- 22 -

85										90										95									
agt	gtc	gca	gtc	gga	aaa	agc	acg	acg	gcg	cgg	atc	ttg	cag	aag	ctg														
Ser	Val	Ala	Val	Gly	Lys	Ser	Thr	Thr	Ala	Arg	Ile	Leu	Gln	Lys	Leu														
			100						105				110																
ctt	tcg	cgt	ttg	cct	gac	cgt	cca	aaa	gtg	agc	ctt	atc	acg	aca	gat														
Leu	Ser	Arg	Leu	Pro	Asp	Arg	Pro	Lys	Val	Ser	Leu	Ile	Thr	Thr	Asp														
		115					120					125																	
ggt	ttt	tta	ttt	cct	act	gcc	gag	ctg	aaa	aag	aaa	aat	atg	atg	tca														
Gly	Phe	Leu	Phe	Pro	Thr	Ala	Glu	Leu	Lys	Lys	Lys	Asn	Met	Met	Ser														
		130				135						140																	
aga	aaa	gga	ttt	cct	gaa	agc	tat	gat	gta	aag	gcg	ctg	ctc	gaa	ttt														
Arg	Lys	Gly	Phe	Pro	Glu	Ser	Tyr	Asp	Val	Lys	Ala	Leu	Leu	Glu	Phe														
					150					155					160														
ttg	aat	gac	tta	aaa	tca	gga	aag	gac	agc	gta	aag	gcc	ccg	gtg	tat														
Leu	Asn	Asp	Leu	Lys	Ser	Gly	Lys	Asp	Ser	Val	Lys	Ala	Pro	Val	Tyr														
				165				170						175															
tcc	cat	cta	acc	tat	gac	cgc	gag	gaa	ggt	gtg	ttc	gag	ggt	gta	gaa														
Ser	His	Leu	Thr	Tyr	Asp	Arg	Glu	Glu	Gly	Val	Phe	Glu	Val	Val	Glu														
			180					185					190																
cag	gcg	gat	att	gtg	att	att	gaa	ggc	att	aat	gtt	ctt	cag	tcg	ccc														
Gln	Ala	Asp	Ile	Val	Ile	Ile	Glu	Gly	Ile	Asn	Val	Leu	Gln	Ser	Pro														
		195					200					205																	
acc	ttg	gag	gat	gac	cgg	gaa	aac	ccg	cgt	att	ttt	gtt	tcc	gat	ttc														
Thr	Leu	Glu	Asp	Asp	Arg	Glu	Asn	Pro	Arg	Ile	Phe	Val	Ser	Asp	Phe														
		210				215					220																		
ttt	gat	ttt	tcg	att	tat	gtg	gat	gcg	gag	gaa	agc	cgg	att	ttc	act														
Phe	Asp	Phe	Ser	Ile	Tyr	Val	Asp	Ala	Glu	Glu	Ser	Arg	Ile	Phe	Thr														
	225				230			235							240														
tgg	tat	tta	gag	cgt	ttt	cgc	ctg	ctt	cgg	gaa	aca	gct	ttt	caa	aat														
Trp	Tyr	Leu	Glu	Arg	Phe	Arg	Leu	Leu	Arg	Glu	Thr	Ala	Phe	Gln	Asn														
				245				250					255																
cct	gat	tca	tat	ttt	cat	aaa	ttt	aaa	gac	ttg	tcc	gat	cag	gag	gct														
Pro	Asp	Ser	Tyr	Phe	His	Lys	Phe	Lys	Asp	Leu	Ser	Asp	Gln	Glu	Ala														
			260				265					270																	
gac	gag	atg	gca	gcc	tcg	att	tgg	gag	agt	gtc	aac	cgg	ccg	aat	tta														
Asp	Glu	Met	Ala	Ala	Ser	Ile	Trp	Glu	Ser	Val	Asn	Arg	Pro	Asn	Leu														
		275					280				285																		
tat	gaa	aat	att	ttg	cca	act	aaa	ttc	agg	tca	gat	ctc	att	ttg	cgt														
Tyr	Glu	Asn	Ile	Leu	Pro	Thr	Lys	Phe	Arg	Ser	Asp	Leu	Ile	Leu	Arg														
	290				295						300																		
aag	gga	gac	ggg	cat	aag	gtc	gag	gaa	gtg	ttg	gta	agg	agg	gta	tga														
Lys	Gly	Asp	Gly	His	Lys	Val	Glu	Glu	Val	Leu	Val	Arg	Arg	Val															
	305				310				315																				

- 23 -

<210> 21  
 <211> 882  
 <212> DNA  
 <213> Bacillus subtilis

<220>  
 <221> CDS  
 <222> (1)..(879)

<400> 21  
 ttg tcg att gct gta tct gaa gaa gag gca aaa gct gtg gaa gga ttg 48  
 Met Ser Ile Ala Val Ser Glu Glu Glu Ala Lys Ala Val Glu Gly Leu  
 1 5 10 15  
 aat gat tat cta tct gtt gaa gaa gtg gag acg atc tat att ccg ctt 96  
 Asn Asp Tyr Leu Ser Val Glu Glu Val Glu Thr Ile Tyr Ile Pro Leu  
 20 25 30  
 gtt cgc ttg ctt cat tta cat gtc aag tct gcg gct gaa cgc aat aag 144  
 Val Arg Leu Leu His Leu His Val Lys Ser Ala Ala Glu Arg Asn Lys  
 35 40 45  
 cat gtc aat gtt ttt ttg aag cac cca cat tca gcc aaa att ccg ttt 192  
 His Val Asn Val Phe Leu Lys His Pro His Ser Ala Lys Ile Pro Phe  
 50 55 60  
 att atc ggc att gcc ggc agt gtc gca gtc gga aaa agc acg acg gcg 240  
 Ile Ile Gly Ile Ala Gly Ser Val Ala Val Gly Lys Ser Thr Thr Ala  
 65 70 75 80  
 cgg atc ttg cag aag ctg ctt tcg cgt ttg cct gac cgt cca aaa gtg 288  
 Arg Ile Leu Gln Lys Leu Leu Ser Arg Leu Pro Asp Arg Pro Lys Val  
 85 90 95  
 agc ctt atc acg aca gat ggt ttt tta ttt cct act gcc gag ctg aaa 336  
 Ser Leu Ile Thr Thr Asp Gly Phe Leu Phe Pro Thr Ala Glu Leu Lys  
 100 105 110  
 aag aaa aat atg atg tca aga aaa gga ttt cct gaa agc tat gat gta 384  
 Lys Lys Asn Met Met Ser Arg Lys Gly Phe Pro Glu Ser Tyr Asp Val  
 115 120 125  
 aag gcg ctg ctc gaa ttt ttg aat gac tta aaa tca gga aag gac agc 432  
 Lys Ala Leu Leu Glu Phe Leu Asn Asp Leu Lys Ser Gly Lys Asp Ser  
 130 135 140  
 gta aag gcc ccg gtg tat tcc cat cta acc tat gac cgc gag gaa ggt 480  
 Val Lys Ala Pro Val Tyr Ser His Leu Thr Tyr Asp Arg Glu Glu Gly  
 145 150 155 160  
 gtg ttc gag gtt gta gaa cag gcg gat att gtg att att gaa ggc att 528  
 Val Phe Glu Val Val Glu Gln Ala Asp Ile Val Ile Ile Glu Gly Ile  
 165 170 175  
 aat gtt ctt cag tcg ccc acc ttg gag gat gac cgg gaa aac ccg cgt 576  
 Asn Val Leu Gln Ser Pro Thr Leu Glu Asp Asp Arg Glu Asn Pro Arg  
 180 185 190

- 24 -

att ttt gtt tcc gat ttc ttt gat ttt tcg att tat gtg gat gcg gag	624
Ile Phe Val Ser Asp Phe Phe Asp Phe Ser Ile Tyr Val Asp Ala Glu	
195 200 205	
gaa agc cgg att ttc act tgg tat tta gag cgt ttt cgc ctg ctt cgg	672
Glu Ser Arg Ile Phe Thr Trp Tyr Leu Glu Arg Phe Arg Leu Leu Arg	
210 215 220	
gaa aca gct ttt caa aat cct gat tca tat ttt cat aaa ttt aaa gac	720
Glu Thr Ala Phe Gln Asn Pro Asp Ser Tyr Phe His Lys Phe Lys Asp	
225 230 235 240	
ttg tcc gat cag gag gct gac gag atg gca gcc tcg att tgg gag agt	768
Leu Ser Asp Gln Glu Ala Asp Glu Met Ala Ala Ser Ile Trp Glu Ser	
245 250 255	
gtc aac cgg ccg aat tta tat gaa aat att ttg cca act aaa ttc agg	816
Val Asn Arg Pro Asn Leu Tyr Glu Asn Ile Leu Pro Thr Lys Phe Arg	
260 265 270	
tca gat ctc att ttg cgt aag gga gac ggg cat aag gtc gag gaa gtg	864
Ser Asp Leu Ile Leu Arg Lys Gly Asp Gly His Lys Val Glu Glu Val	
275 280 285	
ttg gta agg agg gta tga	882
Leu Val Arg Arg Val	
290	

&lt;210&gt; 22

&lt;211&gt; 846

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(843)

&lt;400&gt; 22

gtg gaa gga ttg aat gat tat cta tct gtt gaa gaa gtg gag acg atc	48
Met Glu Gly Leu Asn Asp Tyr Leu Ser Val Glu Glu Val Glu Thr Ile	
1 5 10 15	
tat att ccg ctt gtt cgc ttg ctt cat tta cat gtc aag tct gcg gct	96
Tyr Ile Pro Leu Val Arg Leu Leu His Leu His Val Lys Ser Ala Ala	
20 25 30	
gaa cgc aat aag cat gtc aat gtt ttt ttg aag cac cca cat tca gcc	144
Glu Arg Asn Lys His Val Asn Val Phe Leu Lys His Pro His Ser Ala	
35 40 45	
aaa att ccg ttt att atc ggc att gcc ggc agt gtc gca gtc gga aaa	192
Lys Ile Pro Phe Ile Ile Gly Ile Ala Gly Ser Val Ala Val Gly Lys	
50 55 60	
agc acg acg gcg cgg atc ttg cag aag ctg ctt tcg cgt ttg cct gac	240
Ser Thr Thr Ala Arg Ile Leu Gln Lys Leu Leu Ser Arg Leu Pro Asp	

- 25 -

65					70					75					80	
cgt cca aaa gtg agc ctt atc acg aca gat ggt ttt tta ttt cct act	288															
Arg Pro Lys Val Ser Leu Ile Thr Thr Asp Gly Phe Leu Phe Pro Thr																
85 90 95																
gcc gag ctg aaa aag aaa aat atg atg tca aga aaa gga ttt cct gaa	336															
Ala Glu Leu Lys Lys Lys Asn Met Met Ser Arg Lys Gly Phe Pro Glu																
100 105 110																
agc tat gat gta aag gcg ctg ctc gaa ttt ttg aat gac tta aaa tca	384															
Ser Tyr Asp Val Lys Ala Leu Leu Glu Phe Leu Asn Asp Leu Lys Ser																
115 120 125																
gga aag gac agc gta aag gcc ccg gtg tat tcc cat cta acc tat gac	432															
Gly Lys Asp Ser Val Lys Ala Pro Val Tyr Ser His Leu Thr Tyr Asp																
130 135 140																
cgc gag gaa ggt gtg ttc gag gtt gta gaa cag gcg gat att gtg att	480															
Arg Glu Glu Gly Val Phe Glu Val Val Glu Gln Ala Asp Ile Val Ile																
145 150 155 160																
att gaa ggc att aat gtt ctt cag tcg ccc acc ttg gag gat gac cgg	528															
Ile Glu Gly Ile Asn Val Leu Gln Ser Pro Thr Leu Glu Asp Asp Arg																
165 170 175																
gaa aac ccg cgt att ttt gtt tcc gat ttc ttt gat ttt tcg att tat	576															
Glu Asn Pro Arg Ile Phe Val Ser Asp Phe Phe Asp Phe Ser Ile Tyr																
180 185 190																
gtg gat gcg gag gaa agc ccg att ttc act tgg tat tta gag cgt ttt	624															
Val Asp Ala Glu Glu Ser Arg Ile Phe Thr Trp Tyr Leu Glu Arg Phe																
195 200 205																
cgc ctg ctt ccg gaa aca gct ttt caa aat cct gat tca tat ttt cat	672															
Arg Leu Leu Arg Glu Thr Ala Phe Gln Asn Pro Asp Ser Tyr Phe His																
210 215 220																
aaa ttt aaa gac ttg tcc gat cag gag gct gac gag atg gca gcc tcg	720															
Lys Phe Lys Asp Leu Ser Asp Gln Glu Ala Asp Glu Met Ala Ala Ser																
225 230 235 240																
att tgg gag agt gtc aac ccg ccg aat tta tat gaa aat att ttg cca	768															
Ile Trp Glu Ser Val Asn Arg Pro Asn Leu Tyr Glu Asn Ile Leu Pro																
245 250 255																
act aaa ttc agg tca gat ctc att ttg cgt aag gga gac ggg cat aag	816															
Thr Lys Phe Arg Ser Asp Leu Ile Leu Arg Lys Gly Asp Gly His Lys																
260 265 270																
gtc gag gaa gtg ttg gta agg agg gta tga	846															
Val Glu Glu Val Leu Val Arg Arg Val																
275 280																

&lt;210&gt; 23

&lt;211&gt; 831

&lt;212&gt; DNA



- 26 -

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(831)

&lt;400&gt; 23

atg	aaa	aca	aaa	ctg	gat	ttt	cta	aaa	atg	aag	gag	tct	gaa	gaa	ccg	48
Met	Lys	Thr	Lys	Leu	Asp	Phe	Leu	Lys	Met	Lys	Glu	Ser	Glu	Glu	Pro	
1				5					10					15		
att	gtc	atg	ctg	acc	gct	tat	gat	tat	ccg	gca	gct	aaa	ctt	gct	gaa	96
Ile	Val	Met	Leu	Thr	Ala	Tyr	Asp	Tyr	Pro	Ala	Ala	Lys	Leu	Ala	Glu	
			20					25					30			
caa	gcg	gga	gtt	gac	atg	att	tta	gtc	ggt	gat	tca	ctt	gga	atg	gtc	144
Gln	Ala	Gly	Val	Asp	Met	Ile	Leu	Val	Gly	Asp	Ser	Leu	Gly	Met	Val	
		35					40					45				
gtc	ctc	ggc	ctt	gat	tca	act	gtc	ggt	gtg	aca	gtt	gcg	gac	atg	atc	192
Val	Leu	Gly	Leu	Asp	Ser	Thr	Val	Gly	Val	Thr	Val	Ala	Asp	Met	Ile	
	50					55					60					
cat	cat	aca	aaa	gcc	gtt	aaa	agg	ggt	gcg	ccg	aat	acc	ttt	att	gtg	240
His	His	Thr	Lys	Ala	Val	Lys	Arg	Gly	Ala	Pro	Asn	Thr	Phe	Ile	Val	
65					70					75					80	
aca	gat	atg	ccg	ttt	atg	tct	tat	cac	ctg	tct	aag	gaa	gat	acg	ctg	288
Thr	Asp	Met	Pro	Phe	Met	Ser	Tyr	His	Leu	Ser	Lys	Glu	Asp	Thr	Leu	
				85					90					95		
aaa	aat	gca	gcg	gct	atc	gtt	cag	gaa	agc	gga	gct	gac	gca	ctg	aag	336
Lys	Asn	Ala	Ala	Ala	Ile	Val	Gln	Glu	Ser	Gly	Ala	Asp	Ala	Leu	Lys	
			100					105					110			
ctt	gag	ggc	gga	gaa	ggc	gtg	ttt	gaa	tcc	att	cgc	gca	ttg	acg	ctt	384
Leu	Glu	Gly	Gly	Glu	Gly	Val	Phe	Glu	Ser	Ile	Arg	Ala	Leu	Thr	Leu	
		115					120					125				
gga	ggc	att	cca	gta	gtc	agt	cac	tta	ggt	ttg	aca	ccg	cag	tca	gtc	432
Gly	Gly	Ile	Pro	Val	Val	Ser	His	Leu	Gly	Leu	Thr	Pro	Gln	Ser	Val	
		130				135					140					
ggc	gta	ctg	ggc	ggc	tat	aaa	gta	cag	ggc	aaa	gac	gaa	caa	agc	gcc	480
Gly	Val	Leu	Gly	Gly	Tyr	Lys	Val	Gln	Gly	Lys	Asp	Glu	Gln	Ser	Ala	
145					150					155					160	
aaa	aaa	tta	ata	gaa	gac	agt	ata	aaa	tgc	gaa	gaa	gca	gga	gct	atg	528
Lys	Lys	Leu	Ile	Glu	Asp	Ser	Ile	Lys	Cys	Glu	Glu	Ala	Gly	Ala	Met	
				165					170					175		
atg	ctt	gtg	ctg	gaa	tgt	gtg	ccg	gca	gaa	ctc	aca	gcc	aaa	att	gcc	576
Met	Leu	Val	Leu	Glu	Cys	Val	Pro	Ala	Glu	Leu	Thr	Ala	Lys	Ile	Ala	
			180					185					190			
gag	acg	cta	agc	ata	ccg	gtc	att	gga	atc	ggg	gct	ggt	gtg	aaa	gcg	624
Glu	Thr	Leu	Ser	Ile	Pro	Val	Ile	Gly	Ile	Gly	Ala	Gly	Val	Lys	Ala	
		195				200						205				

- 27 -

gac gga caa gtt ctc gtt tat cat gat att atc ggc cac ggt gtt gag 672  
 Asp Gly Gln Val Leu Val Tyr His Asp Ile Ile Gly His Gly Val Glu  
 210 215 220  
 aga aca cct aaa ttt gta aag caa tat acg cgc att gat gaa acc atc 720  
 Arg Thr Pro Lys Phe Val Lys Gln Tyr Thr Arg Ile Asp Glu Thr Ile  
 225 230 235 240  
 gaa aca gca atc agc gga tat gtt cag gat gta aga cat cgt gct ttc 768  
 Glu Thr Ala Ile Ser Gly Tyr Val Gln Asp Val Arg His Arg Ala Phe  
 245 250 255  
 cct gaa caa aag cat tcc ttt caa atg aac cag aca gtg ctt gac ggc 816  
 Pro Glu Gln Lys His Ser Phe Gln Met Asn Gln Thr Val Leu Asp Gly  
 260 265 270  
 ttg tac ggg gga aaa 831  
 Leu Tyr Gly Gly Lys  
 275

&lt;210&gt; 24

&lt;211&gt; 277

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 24

Met Lys Thr Lys Leu Asp Phe Leu Lys Met Lys Glu Ser Glu Glu Pro  
 1 5 10 15  
 Ile Val Met Leu Thr Ala Tyr Asp Tyr Pro Ala Ala Lys Leu Ala Glu  
 20 25 30  
 Gln Ala Gly Val Asp Met Ile Leu Val Gly Asp Ser Leu Gly Met Val  
 35 40 45  
 Val Leu Gly Leu Asp Ser Thr Val Gly Val Thr Val Ala Asp Met Ile  
 50 55 60  
 His His Thr Lys Ala Val Lys Arg Gly Ala Pro Asn Thr Phe Ile Val  
 65 70 75 80  
 Thr Asp Met Pro Phe Met Ser Tyr His Leu Ser Lys Glu Asp Thr Leu  
 85 90 95  
 Lys Asn Ala Ala Ala Ile Val Gln Glu Ser Gly Ala Asp Ala Leu Lys  
 100 105 110  
 Leu Glu Gly Gly Glu Gly Val Phe Glu Ser Ile Arg Ala Leu Thr Leu  
 115 120 125  
 Gly Gly Ile Pro Val Val Ser His Leu Gly Leu Thr Pro Gln Ser Val  
 130 135 140  
 Gly Val Leu Gly Gly Tyr Lys Val Gln Gly Lys Asp Glu Gln Ser Ala  
 145 150 155 160

- 28 -

Lys Lys Leu Ile Glu Asp Ser Ile Lys Cys Glu Glu Ala Gly Ala Met  
 165 170 175  
 Met Leu Val Leu Glu Cys Val Pro Ala Glu Leu Thr Ala Lys Ile Ala  
 180 185 190  
 Glu Thr Leu Ser Ile Pro Val Ile Gly Ile Gly Ala Gly Val Lys Ala  
 195 200 205  
 Asp Gly Gln Val Leu Val Tyr His Asp Ile Ile Gly His Gly Val Glu  
 210 215 220  
 Arg Thr Pro Lys Phe Val Lys Gln Tyr Thr Arg Ile Asp Glu Thr Ile  
 225 230 235 240  
 Glu Thr Ala Ile Ser Gly Tyr Val Gln Asp Val Arg His Arg Ala Phe  
 245 250 255  
 Pro Glu Gln Lys His Ser Phe Gln Met Asn Gln Thr Val Leu Asp Gly  
 260 265 270  
 Leu Tyr Gly Gly Lys  
 275

<210> 25  
 <211> 858  
 <212> DNA  
 <213> Bacillus subtilis

<220>  
 <221> CDS  
 <222> (1)..(858)

<400> 25  
 atg aga cag att act gat att tca cag ctg aaa gaa gcc ata aaa caa 48  
 Met Arg Gln Ile Thr Asp Ile Ser Gln Leu Lys Glu Ala Ile Lys Gln  
 1 5 10 15  
 tac cat tca gag ggc aag tca atc gga ttt gtt ccg acg atg ggg ttt 96  
 Tyr His Ser Glu Gly Lys Ser Ile Gly Phe Val Pro Thr Met Gly Phe  
 20 25 30  
 ctg cat gag ggg cat tta acc tta gca gac aaa gca aga caa gaa aac 144  
 Leu His Glu Gly His Leu Thr Leu Ala Asp Lys Ala Arg Gln Glu Asn  
 35 40 45  
 gac gcc gtt att atg agt att ttt gtg aat cct gca caa ttc ggc cct 192  
 Asp Ala Val Ile Met Ser Ile Phe Val Asn Pro Ala Gln Phe Gly Pro  
 50 55 60  
 aat gaa gat ttt gaa gca tat ccg cgc gat att gag cgg gat gca gct 240  
 Asn Glu Asp Phe Glu Ala Tyr Pro Arg Asp Ile Glu Arg Asp Ala Ala  
 65 70 75 80  
 ctt gca gaa aac gcc gga gtc gat att ctt ttt acg cca gat gct cat 288  
 Leu Ala Glu Asn Ala Gly Val Asp Ile Leu Phe Thr Pro Asp Ala His  
 85 90 95

- 29 -

gat Asp	atg Met	tat Tyr	ccc Pro 100	ggg Gly	gaa Glu	aag Lys	aat Asn	gtc Val 105	acg Thr	att Ile	cat His	gta Val	gaa Glu 110	aga Arg	cgc Arg	336
aca Thr	gac Asp	gtg Val 115	tta Leu	tgc Cys	ggg Gly	cgc Arg	tca Ser 120	aga Arg	gaa Glu	gga Gly	cat His	ttt Phe 125	gac Asp	ggg Gly	gtc Val	384
gcg Ala	atc Ile 130	gta Val	ctg Leu	acg Thr	aag Lys	ctt Leu 135	ttc Phe	aat Asn	cta Leu	gtc Val	aag Lys 140	ccg Pro	act Thr	cgt Arg	gcc Ala	432
tat Tyr 145	ttc Phe	ggg Gly	tta Leu	aaa Lys	gat Asp 150	gcg Ala	cag Gln	cag Gln	gta Val	gct Ala 155	gtt Val	gtt Val	gat Asp	ggg Gly	tta Leu 160	480
atc Ile	agc Ser	gac Asp	ttc Phe	ttc Phe 165	atg Met	gat Asp	att Ile	gaa Glu 170	ttg Leu	gtt Val	cct Pro	gtc Val	gat Asp	acg Thr 175	gtc Val	528
aga Arg	gag Glu	gaa Glu	gac Asp 180	ggc Gly	tta Leu	gcc Ala	aaa Lys	agc Ser 185	tct Ser	cgc Arg	aat Asn	gta Val	tac Tyr 190	tta Leu	aca Thr	576
gct Ala	gag Glu	gaa Glu 195	aga Arg	aaa Lys	gaa Glu	gcg Ala	cct Pro 200	aag Lys	ctg Leu	tat Tyr	cgg Arg	gcc Ala 205	ctt Leu	caa Gln	aca Thr	624
agt Ser	gcg Ala 210	gaa Glu	ctt Leu	gtc Val	caa Gln	gcc Ala 215	ggg Gly	gaa Glu	aga Arg	gat Asp	cct Pro 220	gaa Glu	gcg Ala	gtg Val	ata Ile	672
aaa Lys 225	gct Ala	gca Ala	aaa Lys	gat Asp	atc Ile 230	att Ile	gaa Glu	acg Thr	act Thr	agc Ser 235	gga Gly	acc Thr	ata Ile	gac Asp	tat Tyr 240	720
gta Val	gag Glu	ctt Leu	tat Tyr	tcc Ser 245	tat Tyr	ccg Pro	gaa Glu	ctc Leu 250	gag Glu	cct Pro	gtg Val	aat Asn	gaa Glu 255	att Ile	gct Ala	768
gga Gly	aag Lys	atg Met	att Ile 260	ctc Leu	gct Ala	gtt Val	gca Ala 265	gtt Val	gct Ala	ttt Phe	tca Ser	aaa Lys 270	gcg Ala	cgt Arg	tta Leu	816
ata Ile	gat Asp	aat Asn 275	atc Ile	att Ile	att Ile	gat Asp	att Ile 280	cga Arg	gaa Glu	atg Met	gag Glu 285	aga Arg	ata Ile			858

<210> 26

<211> 286

<212> PRT

<213> Bacillus subtilis

<400> 26

Met Arg Gln Ile Thr Asp Ile Ser Gln Leu Lys Glu Ala Ile Lys Gln  
1 5 10 15

- 30 -

Tyr His Ser Glu Gly Lys Ser Ile Gly Phe Val Pro Thr Met Gly Phe  
 20 25 30  
 Leu His Glu Gly His Leu Thr Leu Ala Asp Lys Ala Arg Gln Glu Asn  
 35 40 45  
 Asp Ala Val Ile Met Ser Ile Phe Val Asn Pro Ala Gln Phe Gly Pro  
 50 55 60  
 Asn Glu Asp Phe Glu Ala Tyr Pro Arg Asp Ile Glu Arg Asp Ala Ala  
 65 70 75 80  
 Leu Ala Glu Asn Ala Gly Val Asp Ile Leu Phe Thr Pro Asp Ala His  
 85 90 95  
 Asp Met Tyr Pro Gly Glu Lys Asn Val Thr Ile His Val Glu Arg Arg  
 100 105 110  
 Thr Asp Val Leu Cys Gly Arg Ser Arg Glu Gly His Phe Asp Gly Val  
 115 120 125  
 Ala Ile Val Leu Thr Lys Leu Phe Asn Leu Val Lys Pro Thr Arg Ala  
 130 135 140  
 Tyr Phe Gly Leu Lys Asp Ala Gln Gln Val Ala Val Val Asp Gly Leu  
 145 150 155 160  
 Ile Ser Asp Phe Phe Met Asp Ile Glu Leu Val Pro Val Asp Thr Val  
 165 170 175  
 Arg Glu Glu Asp Gly Leu Ala Lys Ser Ser Arg Asn Val Tyr Leu Thr  
 180 185 190  
 Ala Glu Glu Arg Lys Glu Ala Pro Lys Leu Tyr Arg Ala Leu Gln Thr  
 195 200 205  
 Ser Ala Glu Leu Val Gln Ala Gly Glu Arg Asp Pro Glu Ala Val Ile  
 210 215 220  
 Lys Ala Ala Lys Asp Ile Ile Glu Thr Thr Ser Gly Thr Ile Asp Tyr  
 225 230 235 240  
 Val Glu Leu Tyr Ser Tyr Pro Glu Leu Glu Pro Val Asn Glu Ile Ala  
 245 250 255  
 Gly Lys Met Ile Leu Ala Val Ala Val Ala Phe Ser Lys Ala Arg Leu  
 260 265 270  
 Ile Asp Asn Ile Ile Ile Asp Ile Arg Glu Met Glu Arg Ile  
 275 280 285

&lt;210&gt; 27

&lt;211&gt; 381

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

- 31 -

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(381)

&lt;400&gt; 27

atg	tat	cga	aca	atg	atg	agc	ggc	aaa	ctt	cac	agg	gca	act	gtt	acg	48
Met	Tyr	Arg	Thr	Met	Met	Ser	Gly	Lys	Leu	His	Arg	Ala	Thr	Val	Thr	
1				5					10					15		
gaa	gca	aac	ctg	aac	tat	gtg	gga	agc	att	aca	att	gat	gaa	gat	ctc	96
Glu	Ala	Asn	Leu	Asn	Tyr	Val	Gly	Ser	Ile	Thr	Ile	Asp	Glu	Asp	Leu	
			20					25					30			
att	gat	gct	gtg	gga	atg	ctt	cct	aat	gaa	aaa	gta	caa	att	gtg	aat	144
Ile	Asp	Ala	Val	Gly	Met	Leu	Pro	Asn	Glu	Lys	Val	Gln	Ile	Val	Asn	
		35					40					45				
aat	aat	aat	gga	gca	cgt	ctt	gaa	acg	tat	att	att	cct	ggg	aaa	cgg	192
Asn	Asn	Asn	Gly	Ala	Arg	Leu	Glu	Thr	Tyr	Ile	Ile	Pro	Gly	Lys	Arg	
	50					55					60					
gga	agc	ggc	gtc	ata	tgc	tta	aac	ggg	gca	gcc	gca	cgc	ctt	gtg	cag	240
Gly	Ser	Gly	Val	Ile	Cys	Leu	Asn	Gly	Ala	Ala	Ala	Arg	Leu	Val	Gln	
65					70				75						80	
gaa	gga	gat	aag	gtc	att	att	att	tcc	tac	aaa	atg	atg	tct	gat	caa	288
Glu	Gly	Asp	Lys	Val	Ile	Ile	Ile	Ser	Tyr	Lys	Met	Met	Ser	Asp	Gln	
				85					90					95		
gaa	gca	gca	agc	cat	gag	cgc	aaa	gtg	gct	gtt	ctg	aat	gat	caa	aac	336
Glu	Ala	Ala	Ser	His	Glu	Pro	Lys	Val	Ala	Val	Leu	Asn	Asp	Gln	Asn	
			100					105					110			
aaa	att	gaa	caa	atg	ctg	ggg	aac	gaa	cca	gcc	cgt	aca	att	ttg		381
Lys	Ile	Glu	Gln	Met	Leu	Gly	Asn	Glu	Pro	Ala	Arg	Thr	Ile	Leu		
		115					120					125				

&lt;210&gt; 28

&lt;211&gt; 127

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 28

Met	Tyr	Arg	Thr	Met	Met	Ser	Gly	Lys	Leu	His	Arg	Ala	Thr	Val	Thr	
1				5					10					15		
Glu	Ala	Asn	Leu	Asn	Tyr	Val	Gly	Ser	Ile	Thr	Ile	Asp	Glu	Asp	Leu	
			20					25					30			
Ile	Asp	Ala	Val	Gly	Met	Leu	Pro	Asn	Glu	Lys	Val	Gln	Ile	Val	Asn	
		35					40					45				
Asn	Asn	Asn	Gly	Ala	Arg	Leu	Glu	Thr	Tyr	Ile	Ile	Pro	Gly	Lys	Arg	
	50					55					60					
Gly	Ser	Gly	Val	Ile	Cys	Leu	Asn	Gly	Ala	Ala	Ala	Arg	Leu	Val	Gln	
65					70				75						80	

- 32 -

Glu Gly Asp Lys Val Ile Ile Ile Ser Tyr Lys Met Met Ser Asp Gln  
85 90 95

Glu Ala Ala Ser His Glu Pro Lys Val Ala Val Leu Asn Asp Gln Asn  
100 105 110

Lys Ile Glu Gln Met Leu Gly Asn Glu Pro Ala Arg Thr Ile Leu  
115 120 125

<210> 29

<211> 894

<212> DNA

<213> Bacillus subtilis

<220>

<221> CDS

<222> (1)..(894)

<400> 29

atg aaa att gga att atc ggc gga ggc tcc gtt ggt ctt tta tgc gcc 48  
Met Lys Ile Gly Ile Ile Gly Gly Gly Ser Val Gly Leu Leu Cys Ala  
1 5 10 15

tat tat ttg tca ctt tat cac gac gtg act gtt gtg acg agg cgg caa 96  
Tyr Tyr Leu Ser Leu Tyr His Asp Val Thr Val Val Thr Arg Arg Gln  
20 25 30

gaa cag gct gcg gcc att cag tct gaa gga atc cgg ctt tat aaa ggc 144  
Glu Gln Ala Ala Ala Ile Gln Ser Glu Gly Ile Arg Leu Tyr Lys Gly  
35 40 45

ggg gag gaa ttc agg gct gat tgc agt gcg gac acg agt atc aat tcg 192  
Gly Glu Glu Phe Arg Ala Asp Cys Ser Ala Asp Thr Ser Ile Asn Ser  
50 55 60

gac ttt gac ctg ctt gtc gtg aca gtg aag cag cat cag ctt caa tct 240  
Asp Phe Asp Leu Leu Val Val Thr Val Lys Gln His Gln Leu Gln Ser  
65 70 75 80

gtt ttt tcg tcg ctt gaa cga atc ggg aag acg aat ata tta ttt ttg 288  
Val Phe Ser Ser Leu Glu Arg Ile Gly Lys Thr Asn Ile Leu Phe Leu  
85 90 95

caa aac ggc atg ggg cat atc cac gac cta aaa gac tgg cac gtt ggc 336  
Gln Asn Gly Met Gly His Ile His Asp Leu Lys Asp Trp His Val Gly  
100 105 110

cat tcc att tat gtt gga atc gtt gag cac gga gct gta aga aaa tcg 384  
His Ser Ile Tyr Val Gly Ile Val Glu His Gly Ala Val Arg Lys Ser  
115 120 125

gat aca gct gtt gat cat aca ggc cta ggt gcg ata aaa tgg agc gcg 432  
Asp Thr Ala Val Asp His Thr Gly Leu Gly Ala Ile Lys Trp Ser Ala  
130 135 140

ttc gac gat gct gaa cca gac cgg ctg aac atc ttg ttt cag cat aac 480

- 33 -

Phe	Asp	Asp	Ala	Glu	Pro	Asp	Arg	Leu	Asn	Ile	Leu	Phe	Gln	His	Asn		
145					150					155					160		
cat	tcg	gat	ttt	ccg	att	tat	tat	gag	acg	gat	tgg	tac	cgt	ctg	ctg	528	
His	Ser	Asp	Phe	Pro	Ile	Tyr	Tyr	Glu	Thr	Asp	Trp	Tyr	Arg	Leu	Leu		
				165					170					175			
acg	ggc	aag	ctg	att	gta	aat	gcg	tgt	att	aat	cct	tta	act	gcg	tta	576	
Thr	Gly	Lys	Leu	Ile	Val	Asn	Ala	Cys	Ile	Asn	Pro	Leu	Thr	Ala	Leu		
			180					185					190				
ttg	caa	gtg	aaa	aat	gga	gaa	ctg	ctg	aca	acg	cca	gct	tat	ctg	gct	624	
Leu	Gln	Val	Lys	Asn	Gly	Glu	Leu	Leu	Thr	Thr	Pro	Ala	Tyr	Leu	Ala		
			195				200					205					
ttt	atg	aag	ctg	gta	ttt	cag	gag	gca	tgc	cgc	att	tta	aaa	ctt	gaa	672	
Phe	Met	Lys	Leu	Val	Phe	Gln	Glu	Ala	Cys	Arg	Ile	Leu	Lys	Leu	Glu		
	210					215					220						
aat	gaa	gaa	aag	gct	tgg	gag	cgg	gtt	cag	gcc	gtt	tgt	ggg	caa	acg	720	
Asn	Glu	Glu	Lys	Ala	Trp	Glu	Arg	Val	Gln	Ala	Val	Cys	Gly	Gln	Thr		
	225				230					235					240		
aaa	gag	aat	cgt	tca	tca	atg	ctg	gtt	gac	gtc	att	gga	ggc	cgg	cag	768	
Lys	Glu	Asn	Arg	Ser	Ser	Met	Leu	Val	Asp	Val	Ile	Gly	Gly	Arg	Gln		
				245					250					255			
acg	gaa	gct	gac	gcc	att	atc	gga	tac	tta	ttg	aag	gaa	gca	agt	ctt	816	
Thr	Glu	Ala	Asp	Ala	Ile	Ile	Gly	Tyr	Leu	Leu	Lys	Glu	Ala	Ser	Leu		
			260					265					270				
caa	ggt	ctt	gat	gcc	gtc	cac	cta	gag	ttt	tta	tat	ggc	agc	atc	aaa	864	
Gln	Gly	Leu	Asp	Ala	Val	His	Leu	Glu	Phe	Leu	Tyr	Gly	Ser	Ile	Lys		
		275					280					285					
gca	ttg	gag	cga	aat	aca	aac	aaa	gtc	ttt							894	
Ala	Leu	Glu	Arg	Asn	Thr	Asn	Lys	Val	Phe								
	290					295											

&lt;210&gt; 30

&lt;211&gt; 298

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 30

Met	Lys	Ile	Gly	Ile	Ile	Gly	Gly	Gly	Ser	Val	Gly	Leu	Leu	Cys	Ala
1				5					10					15	

Tyr	Tyr	Leu	Ser	Leu	Tyr	His	Asp	Val	Thr	Val	Val	Thr	Arg	Arg	Gln
			20					25					30		

Glu	Gln	Ala	Ala	Ala	Ile	Gln	Ser	Glu	Gly	Ile	Arg	Leu	Tyr	Lys	Gly
		35					40					45			

Gly	Glu	Glu	Phe	Arg	Ala	Asp	Cys	Ser	Ala	Asp	Thr	Ser	Ile	Asn	Ser
	50					55					60				



- 34 -

Asp Phe Asp Leu Leu Val Val Thr Val Lys Gln His Gln Leu Gln Ser  
 65 70 75 80  
 Val Phe Ser Ser Leu Glu Arg Ile Gly Lys Thr Asn Ile Leu Phe Leu  
 85 90 95  
 Gln Asn Gly Met Gly His Ile His Asp Leu Lys Asp Trp His Val Gly  
 100 105 110  
 His Ser Ile Tyr Val Gly Ile Val Glu His Gly Ala Val Arg Lys Ser  
 115 120 125  
 Asp Thr Ala Val Asp His Thr Gly Leu Gly Ala Ile Lys Trp Ser Ala  
 130 135 140  
 Phe Asp Asp Ala Glu Pro Asp Arg Leu Asn Ile Leu Phe Gln His Asn  
 145 150 155 160  
 His Ser Asp Phe Pro Ile Tyr Tyr Glu Thr Asp Trp Tyr Arg Leu Leu  
 165 170 175  
 Thr Gly Lys Leu Ile Val Asn Ala Cys Ile Asn Pro Leu Thr Ala Leu  
 180 185 190  
 Leu Gln Val Lys Asn Gly Glu Leu Leu Thr Thr Pro Ala Tyr Leu Ala  
 195 200 205  
 Phe Met Lys Leu Val Phe Gln Glu Ala Cys Arg Ile Leu Lys Leu Glu  
 210 215 220  
 Asn Glu Glu Lys Ala Trp Glu Arg Val Gln Ala Val Cys Gly Gln Thr  
 225 230 235 240  
 Lys Glu Asn Arg Ser Ser Met Leu Val Asp Val Ile Gly Gly Arg Gln  
 245 250 255  
 Thr Glu Ala Asp Ala Ile Ile Gly Tyr Leu Leu Lys Glu Ala Ser Leu  
 260 265 270  
 Gln Gly Leu Asp Ala Val His Leu Glu Phe Leu Tyr Gly Ser Ile Lys  
 275 280 285  
 Ala Leu Glu Arg Asn Thr Asn Lys Val Phe  
 290 295

&lt;210&gt; 31

&lt;211&gt; 1725

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1722)

&lt;400&gt; 31

atg ggg act aat gta cag gtg gat tca gca tct gcc gaa tgt aca cag 48  
 Met Gly Thr Asn Val Gln Val Asp Ser Ala Ser Ala Glu Cys Thr Gln

- 35 -

1					5					10					15	
acg	atg	agc	gga	gca	tta	atg	ctg	att	gaa	tca	tta	aaa	aaa	gag	aaa	96
Thr	Met	Ser	Gly	Ala	Leu	Met	Leu	Ile	Glu	Ser	Leu	Lys	Lys	Glu	Lys	
			20				25						30			
gta	gaa	atg	atc	ttc	ggt	tat	ccg	ggc	ggg	gct	gtg	ctt	ccg	att	tac	144
Val	Glu	Met	Ile	Phe	Gly	Tyr	Pro	Gly	Gly	Ala	Val	Leu	Pro	Ile	Tyr	
			35				40						45			
gat	aag	cta	tac	aat	tca	ggg	ttg	gta	cat	atc	ctt	ccc	cgt	cac	gaa	192
Asp	Lys	Leu	Tyr	Asn	Ser	Gly	Leu	Val	His	Ile	Leu	Pro	Arg	His	Glu	
			50				55						60			
caa	gga	gca	att	cat	gca	gcg	gag	gga	tac	gca	agg	gtc	tcc	gga	aaa	240
Gln	Gly	Ala	Ile	His	Ala	Ala	Glu	Gly	Tyr	Ala	Arg	Val	Ser	Gly	Lys	
			65				70						75			80
ccg	ggt	gtc	gtc	att	gcc	acg	tca	ggg	ccg	gga	gcg	aca	aac	ctt	gtt	288
Pro	Gly	Val	Val	Ile	Ala	Thr	Ser	Gly	Pro	Gly	Ala	Thr	Asn	Leu	Val	
			85						90						95	
aca	ggc	ctt	gct	gat	gcc	atg	att	gat	tca	ttg	ccg	tta	gtc	gtc	ttt	336
Thr	Gly	Leu	Ala	Asp	Ala	Met	Ile	Asp	Ser	Leu	Pro	Leu	Val	Val	Phe	
			100						105						110	
aca	ggg	cag	gta	gca	acc	tct	gta	atc	ggg	agc	gat	gca	ttt	cag	gaa	384
Thr	Gly	Gln	Val	Ala	Thr	Ser	Val	Ile	Gly	Ser	Asp	Ala	Phe	Gln	Glu	
			115			120						125				
gca	gac	att	tta	ggg	att	acg	atg	cca	gta	aca	aaa	cac	agc	tac	cag	432
Ala	Asp	Ile	Leu	Gly	Ile	Thr	Met	Pro	Val	Thr	Lys	His	Ser	Tyr	Gln	
			130			135						140				
gtt	cgc	cag	ccg	gaa	gat	ctg	ccg	cgc	atc	att	aaa	gaa	gcg	ttc	cat	480
Val	Arg	Gln	Pro	Glu	Asp	Leu	Pro	Arg	Ile	Ile	Lys	Glu	Ala	Phe	His	
145			150						155						160	
att	gca	aca	act	gga	aga	ccc	gga	cct	gta	ttg	att	gat	att	ccg	aaa	528
Ile	Ala	Thr	Thr	Gly	Arg	Pro	Gly	Pro	Val	Leu	Ile	Asp	Ile	Pro	Lys	
			165						170						175	
gat	gta	gca	aca	att	gaa	gga	gaa	ttc	agc	tac	gat	cat	gag	atg	aat	576
Asp	Val	Ala	Thr	Ile	Glu	Gly	Glu	Phe	Ser	Tyr	Asp	His	Glu	Met	Asn	
			180			185						190				
ctc	ccg	gga	tac	cag	ccg	aca	aca	gag	ccg	aat	tat	ttg	cag	atc	cgc	624
Leu	Pro	Gly	Tyr	Gln	Pro	Thr	Thr	Glu	Pro	Asn	Tyr	Leu	Gln	Ile	Arg	
			195			200						205				
aag	ctt	gtg	gaa	gcc	gtg	agc	agt	gcg	aaa	aaa	ccg	gtg	atc	ctg	gcg	672
Lys	Leu	Val	Glu	Ala	Val	Ser	Ser	Ala	Lys	Lys	Pro	Val	Ile	Leu	Ala	
			210			215						220				
ggt	gcg	ggc	gta	ctg	cac	gga	aaa	gcg	tca	gaa	gaa	tta	aaa	aat	tat	720
Gly	Ala	Gly	Val	Leu	His	Gly	Lys	Ala	Ser	Glu	Glu	Leu	Lys	Asn	Tyr	
225			230						235						240	

- 36 -

gct gaa cag cag caa atc cct gtg gca cac acc ctt ttg ggg ctc gga	768
Ala Glu Gln Gln Gln Ile Pro Val Ala His Thr Leu Leu Gly Leu Gly	
245 250 255	
ggc ttc ccg gct gac cat ccg ctt ttc cta ggg atg gcg gga atg cac	816
Gly Phe Pro Ala Asp His Pro Leu Phe Leu Gly Met Ala Gly Met His	
260 265 270	
ggt act tat aca gcc aat atg gcc ctt cat gaa tgt gat cta tta atc	864
Gly Thr Tyr Thr Ala Asn Met Ala Leu His Glu Cys Asp Leu Leu Ile	
275 280 285	
agt atc ggc gcc cgt ttt gat gac cgt gtc aca gga aac ctg aaa cac	912
Ser Ile Gly Ala Arg Phe Asp Asp Arg Val Thr Gly Asn Leu Lys His	
290 295 300	
ttt gcc aga aac gca aag ata gcc cac atc gat att gat cca gct gaa	960
Phe Ala Arg Asn Ala Lys Ile Ala His Ile Asp Ile Asp Pro Ala Glu	
305 310 315 320	
atc gga aaa atc atg aaa aca cag att cct gta gtc gga gac agc aaa	1008
Ile Gly Lys Ile Met Lys Thr Gln Ile Pro Val Val Gly Asp Ser Lys	
325 330 335	
att gtc ctg cag gag ctg atc aaa caa gac ggc aaa caa agc gat tca	1056
Ile Val Leu Gln Glu Leu Ile Lys Gln Asp Gly Lys Gln Ser Asp Ser	
340 345 350	
agc gaa tgg aaa aaa cag ctc gca gaa tgg aaa gaa gag tat ccg ctc	1104
Ser Glu Trp Lys Lys Gln Leu Ala Glu Trp Lys Glu Glu Tyr Pro Leu	
355 360 365	
tgg tat gta gat aat gaa gaa gaa ggt ttt aaa cct cag aaa ttg att	1152
Trp Tyr Val Asp Asn Glu Glu Glu Gly Phe Lys Pro Gln Lys Leu Ile	
370 375 380	
gaa tat att cat caa ttt aca aaa gga gag gcc att gtc gca acg gat	1200
Glu Tyr Ile His Gln Phe Thr Lys Gly Glu Ala Ile Val Ala Thr Asp	
385 390 395 400	
gta ggc cag cat caa atg tgg tca gcg caa ttt tat ccg ttc caa aaa	1248
Val Gly Gln His Gln Met Trp Ser Ala Gln Phe Tyr Pro Phe Gln Lys	
405 410 415	
gca gat aaa tgg gtc acg tca ggc gga ctt gga acg atg gga ttc ggt	1296
Ala Asp Lys Trp Val Thr Ser Gly Gly Leu Gly Thr Met Gly Phe Gly	
420 425 430	
ctt ccg gcg gcg atc ggc gca cag ctg gcc gaa aaa gat gct act gtt	1344
Leu Pro Ala Ala Ile Gly Ala Gln Leu Ala Glu Lys Asp Ala Thr Val	
435 440 445	
gtc gcg gtt gtc gga gac ggc gga ttc caa atg acg ctt caa gaa ctc	1392
Val Ala Val Val Gly Asp Gly Gly Phe Gln Met Thr Leu Gln Glu Leu	
450 455 460	
gat gtt att cgc gaa tta aat ctt ccg gtc aag gta gtg att tta aat	1440
Asp Val Ile Arg Glu Leu Asn Leu Pro Val Lys Val Val Ile Leu Asn	

- 37 -

465	470	475	480	
aac gct tgt ctc gga atg gtc aga cag tgg cag gaa att ttc tat gaa				1488
Asn Ala Cys Leu Gly Met Val Arg Gln Trp Gln Glu Ile Phe Tyr Glu				
485		490	495	
gaa cgt tat tca gaa tct aaa ttc gct tct cag cct gac ttc gtc aaa				1536
Glu Arg Tyr Ser Glu Ser Lys Phe Ala Ser Gln Pro Asp Phe Val Lys				
500		505	510	
ttg tcc gaa gca tac ggc att aaa ggc atc aga att tca tca gaa gcg				1584
Leu Ser Glu Ala Tyr Gly Ile Lys Gly Ile Arg Ile Ser Ser Glu Ala				
515		520	525	
gaa gca aag gaa aag ctg gaa gag gca tta aca tca aga gaa cct gtt				1632
Glu Ala Lys Glu Lys Leu Glu Glu Ala Leu Thr Ser Arg Glu Pro Val				
530	535	540		
gtc att gac gtg cgg gtt gcc agc gaa gaa aaa gta ttc ccg atg gtg				1680
Val Ile Asp Val Arg Val Ala Ser Glu Glu Lys Val Phe Pro Met Val				
545	550	555	560	
gct ccg ggg aaa ggg ctg cat gaa atg gtg ggg gtg aaa cct tga				1725
Ala Pro Gly Lys Gly Leu His Glu Met Val Gly Val Lys Pro				
565	570			

&lt;210&gt; 32

&lt;211&gt; 574

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 32

Met Gly Thr Asn Val Gln Val Asp Ser Ala Ser Ala Glu Cys Thr Gln
1 5 10 15

Thr Met Ser Gly Ala Leu Met Leu Ile Glu Ser Leu Lys Lys Glu Lys
20 25 30

Val Glu Met Ile Phe Gly Tyr Pro Gly Gly Ala Val Leu Pro Ile Tyr
35 40 45

Asp Lys Leu Tyr Asn Ser Gly Leu Val His Ile Leu Pro Arg His Glu
50 55 60

Gln Gly Ala Ile His Ala Ala Glu Gly Tyr Ala Arg Val Ser Gly Lys
65 70 75 80

Pro Gly Val Val Ile Ala Thr Ser Gly Pro Gly Ala Thr Asn Leu Val
85 90 95

Thr Gly Leu Ala Asp Ala Met Ile Asp Ser Leu Pro Leu Val Val Phe
100 105 110

Thr Gly Gln Val Ala Thr Ser Val Ile Gly Ser Asp Ala Phe Gln Glu
115 120 125

Ala Asp Ile Leu Gly Ile Thr Met Pro Val Thr Lys His Ser Tyr Gln
---

- 38 -

130		135		140											
Val	Arg	Gln	Pro	Glu	Asp	Leu	Pro	Arg	Ile	Ile	Lys	Glu	Ala	Phe	His
145					150					155					160
Ile	Ala	Thr	Thr	Gly	Arg	Pro	Gly	Pro	Val	Leu	Ile	Asp	Ile	Pro	Lys
				165					170					175	
Asp	Val	Ala	Thr	Ile	Glu	Gly	Glu	Phe	Ser	Tyr	Asp	His	Glu	Met	Asn
			180					185					190		
Leu	Pro	Gly	Tyr	Gln	Pro	Thr	Thr	Glu	Pro	Asn	Tyr	Leu	Gln	Ile	Arg
		195					200					205			
Lys	Leu	Val	Glu	Ala	Val	Ser	Ser	Ala	Lys	Lys	Pro	Val	Ile	Leu	Ala
	210					215					220				
Gly	Ala	Gly	Val	Leu	His	Gly	Lys	Ala	Ser	Glu	Glu	Leu	Lys	Asn	Tyr
225					230					235					240
Ala	Glu	Gln	Gln	Gln	Ile	Pro	Val	Ala	His	Thr	Leu	Leu	Gly	Leu	Gly
				245					250					255	
Gly	Phe	Pro	Ala	Asp	His	Pro	Leu	Phe	Leu	Gly	Met	Ala	Gly	Met	His
			260					265					270		
Gly	Thr	Tyr	Thr	Ala	Asn	Met	Ala	Leu	His	Glu	Cys	Asp	Leu	Leu	Ile
		275					280					285			
Ser	Ile	Gly	Ala	Arg	Phe	Asp	Asp	Arg	Val	Thr	Gly	Asn	Leu	Lys	His
	290					295					300				
Phe	Ala	Arg	Asn	Ala	Lys	Ile	Ala	His	Ile	Asp	Ile	Asp	Pro	Ala	Glu
305					310					315					320
Ile	Gly	Lys	Ile	Met	Lys	Thr	Gln	Ile	Pro	Val	Val	Gly	Asp	Ser	Lys
				325					330					335	
Ile	Val	Leu	Gln	Glu	Leu	Ile	Lys	Gln	Asp	Gly	Lys	Gln	Ser	Asp	Ser
			340					345					350		
Ser	Glu	Trp	Lys	Lys	Gln	Leu	Ala	Glu	Trp	Lys	Glu	Glu	Tyr	Pro	Leu
		355					360					365			
Trp	Tyr	Val	Asp	Asn	Glu	Glu	Glu	Gly	Phe	Lys	Pro	Gln	Lys	Leu	Ile
	370					375					380				
Glu	Tyr	Ile	His	Gln	Phe	Thr	Lys	Gly	Glu	Ala	Ile	Val	Ala	Thr	Asp
385					390					395					400
Val	Gly	Gln	His	Gln	Met	Trp	Ser	Ala	Gln	Phe	Tyr	Pro	Phe	Gln	Lys
				405					410					415	
Ala	Asp	Lys	Trp	Val	Thr	Ser	Gly	Gly	Leu	Gly	Thr	Met	Gly	Phe	Gly
			420					425					430		
Leu	Pro	Ala	Ala	Ile	Gly	Ala	Gln	Leu	Ala	Glu	Lys	Asp	Ala	Thr	Val
		435					440					445			

- 39 -

Val Ala Val Val Gly Asp Gly Gly Phe Gln Met Thr Leu Gln Glu Leu  
 450 455 460

Asp Val Ile Arg Glu Leu Asn Leu Pro Val Lys Val Val Ile Leu Asn  
 465 470 475 480

Asn Ala Cys Leu Gly Met Val Arg Gln Trp Gln Glu Ile Phe Tyr Glu  
 485 490 495

Glu Arg Tyr Ser Glu Ser Lys Phe Ala Ser Gln Pro Asp Phe Val Lys  
 500 505 510

Leu Ser Glu Ala Tyr Gly Ile Lys Gly Ile Arg Ile Ser Ser Glu Ala  
 515 520 525

Glu Ala Lys Glu Lys Leu Glu Glu Ala Leu Thr Ser Arg Glu Pro Val  
 530 535 540

Val Ile Asp Val Arg Val Ala Ser Glu Glu Lys Val Phe Pro Met Val  
 545 550 555 560

Ala Pro Gly Lys Gly Leu His Glu Met Val Gly Val Lys Pro  
 565 570

&lt;210&gt; 33

&lt;211&gt; 525

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(522)

&lt;400&gt; 33

ttg aaa aga att atc aca ttg act gtg gtg aac cgc tcc ggg gtg tta 48  
 Met Lys Arg Ile Ile Thr Leu Thr Val Val Asn Arg Ser Gly Val Leu  
 1 5 10 15

aac cgg atc acc ggt cta ttc aca aaa agg cat tac aac att gaa agc 96  
 Asn Arg Ile Thr Gly Leu Phe Thr Lys Arg His Tyr Asn Ile Glu Ser  
 20 25 30

att aca gtt gga cac aca gaa aca gcc ggc gtt tcc aga atc acc ttc 144  
 Ile Thr Val Gly His Thr Glu Thr Ala Gly Val Ser Arg Ile Thr Phe  
 35 40 45

gtc gtt cat gtt gaa ggt gaa aat gat gtt gaa cag tta acg aaa cag 192  
 Val Val His Val Glu Gly Glu Asn Asp Val Glu Gln Leu Thr Lys Gln  
 50 55 60

ctc aac aaa cag att gat gtg ctg aaa gtc aca gac atc aca aat caa 240  
 Leu Asn Lys Gln Ile Asp Val Leu Lys Val Thr Asp Ile Thr Asn Gln  
 65 70 75 80

tcg att gtc cag agg gag ctg gcc tta atc aag gtt gtc tcc gca cct 288  
 Ser Ile Val Gln Arg Glu Leu Ala Leu Ile Lys Val Val Ser Ala Pro

- 40 -

85										90					95					
tca	aca	aga	aca	gag	att	aat	gga	atc	ata	gaa	ccg	ttt	aga	gcc	tct	336				
Ser	Thr	Arg	Thr	Glu	Ile	Asn	Gly	Ile	Ile	Glu	Pro	Phe	Arg	Ala	Ser					
			100					105					110							
gtc	gtt	gat	gtc	agc	aga	gac	agc	atc	gtt	gtt	cag	gtg	aca	ggg	gaa	384				
Val	Val	Asp	Val	Ser	Arg	Asp	Ser	Ile	Val	Val	Gln	Val	Thr	Gly	Glu					
		115					120					125								
tct	aac	aaa	att	gaa	gcg	ctt	att	gag	tta	tta	aaa	cct	tat	ggc	att	432				
Ser	Asn	Lys	Ile	Glu	Ala	Leu	Ile	Glu	Leu	Leu	Lys	Pro	Tyr	Gly	Ile					
	130					135					140									
aaa	gaa	atc	gcg	aga	aca	ggg	aca	acg	gct	ttt	gcg	agg	gga	acc	agc	480				
Lys	Glu	Ile	Ala	Arg	Thr	Gly	Thr	Thr	Ala	Phe	Ala	Arg	Gly	Thr	Ser					
	145				150					155					160					
aaa	agg	cgt	cat	cca	ata	aaa	caa	tat	cta	ttg	tat	aaa	aca	taa		525				
Lys	Arg	Arg	His	Pro	Ile	Lys	Gln	Tyr	Leu	Leu	Tyr	Lys	Thr							
			165					170												

&lt;210&gt; 34

&lt;211&gt; 174

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 34

Met	Lys	Arg	Ile	Ile	Thr	Leu	Thr	Val	Val	Asn	Arg	Ser	Gly	Val	Leu
1				5					10					15	
Asn	Arg	Ile	Thr	Gly	Leu	Phe	Thr	Lys	Arg	His	Tyr	Asn	Ile	Glu	Ser
			20					25					30		
Ile	Thr	Val	Gly	His	Thr	Glu	Thr	Ala	Gly	Val	Ser	Arg	Ile	Thr	Phe
		35					40					45			
Val	Val	His	Val	Glu	Gly	Glu	Asn	Asp	Val	Glu	Gln	Leu	Thr	Lys	Gln
	50				55						60				
Leu	Asn	Lys	Gln	Ile	Asp	Val	Leu	Lys	Val	Thr	Asp	Ile	Thr	Asn	Gln
	65				70					75				80	
Ser	Ile	Val	Gln	Arg	Glu	Leu	Ala	Leu	Ile	Lys	Val	Val	Ser	Ala	Pro
			85						90				95		
Ser	Thr	Arg	Thr	Glu	Ile	Asn	Gly	Ile	Ile	Glu	Pro	Phe	Arg	Ala	Ser
			100					105					110		
Val	Val	Asp	Val	Ser	Arg	Asp	Ser	Ile	Val	Val	Gln	Val	Thr	Gly	Glu
		115					120					125			
Ser	Asn	Lys	Ile	Glu	Ala	Leu	Ile	Glu	Leu	Leu	Lys	Pro	Tyr	Gly	Ile
	130					135					140				
Lys	Glu	Ile	Ala	Arg	Thr	Gly	Thr	Thr	Ala	Phe	Ala	Arg	Gly	Thr	Ser
	145				150					155					160

- 41 -

Lys Arg Arg His Pro Ile Lys Gln Tyr Leu Leu Tyr Lys Thr  
 165 170

&lt;210&gt; 35

&lt;211&gt; 1029

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; cbs

&lt;222&gt; (1)..(1026)

&lt;400&gt; 35

atg gta aaa gta tat tat aac ggt gat atc aaa gag aac gta ttg gct 48  
 Met Val Lys Val Tyr Tyr Asn Gly Asp Ile Lys Glu Asn Val Leu Ala  
 1 5 10 15

gga aaa aca gta gcg gtt atc ggg tac ggt tcg caa ggc cac gca cat 96  
 Gly Lys Thr Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His  
 20 25 30

gcc ctg aac ctt aaa gaa agc gga gta gac gtg atc gtc ggt gtt aga 144  
 Ala Leu Asn Leu Lys Glu Ser Gly Val Asp Val Ile Val Gly Val Arg  
 35 40 45

caa gga aaa tct ttc act caa gcc caa gaa gac gga cat aaa gta ttt 192  
 Gln Gly Lys Ser Phe Thr Gln Ala Gln Glu Asp Gly His Lys Val Phe  
 50 55 60

tca gta aaa gaa gcg gca gcc caa gcc gaa atc atc atg gtt ctg ctt 240  
 Ser Val Lys Glu Ala Ala Gln Ala Glu Ile Ile Met Val Leu Leu  
 65 70 75 80

ccg gat gag cag cag caa aaa gta tac gaa gct gaa atc aaa gat gaa 288  
 Pro Asp Glu Gln Gln Gln Lys Val Tyr Glu Ala Glu Ile Lys Asp Glu  
 85 90 95

ttg aca gca gga aaa tca tta gta ttc gct cat gga ttt aac gtg cat 336  
 Leu Thr Ala Gly Lys Ser Leu Val Phe Ala His Gly Phe Asn Val His  
 100 105 110

ttc cat caa att gtt cct ccg gcg gat gta gat gta ttc tta gtg gcc 384  
 Phe His Gln Ile Val Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala  
 115 120 125

cct aaa ggc ccg gga cac ttg gta aga aga aca tat gag caa gga gct 432  
 Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Glu Gln Gly Ala  
 130 135 140

ggc gta cct gca ttg ttc gca atc tat caa gat gtg act gga gaa gca 480  
 Gly Val Pro Ala Leu Phe Ala Ile Tyr Gln Asp Val Thr Gly Glu Ala  
 145 150 155 160

aga gac aaa gcc ctc gct tat gct aaa gga atc ggc ggc gca aga gcg 528  
 Arg Asp Lys Ala Leu Ala Tyr Ala Lys Gly Ile Gly Gly Ala Arg Ala  
 165 170 175



- 42 -

ggc gta tta gaa acg aca ttt aaa gaa gaa aca gaa aca gat ttg ttc 576  
 Gly Val Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe  
 180 185 190

ggt gag caa gca gtt ctt tgc ggc gga tta agc gcg ctt gtc aaa gcc 624  
 Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Ser Ala Leu Val Lys Ala  
 195 200 205

gga ttt gaa acc tta act gaa gca ggt tat cag cct gaa ctt gca tac 672  
 Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr  
 210 215 220

ttc gag tgt ctt cat gag ctg aaa tta atc gta gac ctt atg tac gaa 720  
 Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu  
 225 230 235 240

gaa gga ctt gca gga atg aga tat tca atc tct gac aca gca cag tgg 768  
 Glu Gly Leu Ala Gly Met Arg Tyr Ser Ile Ser Asp Thr Ala Gln Trp  
 245 250 255

gga gat ttc gta tca ggc cct cgc gtt gtg gac gcc aaa gta aaa gaa 816  
 Gly Asp Phe Val Ser Gly Pro Arg Val Val Asp Ala Lys Val Lys Glu  
 260 265 270

tct atg aaa gaa gta tta aaa gat atc caa aac ggt aca ttc gca aaa 864  
 Ser Met Lys Glu Val Leu Lys Asp Ile Gln Asn Gly Thr Phe Ala Lys  
 275 280 285

gag tgg atc gtc gaa aac caa gta aac cgt cct cgt ttc aac gct atc 912  
 Glu Trp Ile Val Glu Asn Gln Val Asn Arg Pro Arg Phe Asn Ala Ile  
 290 295 300

aat gca agc gag aac gaa cat caa atc gaa gta gtg gga aga aag ctt 960  
 Asn Ala Ser Glu Asn Glu His Gln Ile Glu Val Val Gly Arg Lys Leu  
 305 310 315 320

cgt gaa atg atg ccg ttt gtg aaa caa ggc aag aag aag gaa gcg gtg 1008  
 Arg Glu Met Met Pro Phe Val Lys Gln Gly Lys Lys Lys Glu Ala Val  
 325 330 335

gtc tcc gtt gcg caa aat taa 1029  
 Val Ser Val Ala Gln Asn  
 340

&lt;210&gt; 36

&lt;211&gt; 342

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 36

Met Val Lys Val Tyr Tyr Asn Gly Asp Ile Lys Glu Asn Val Leu Ala  
 1 5 10 15

Gly Lys Thr Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His  
 20 25 30

- 43 -

Ala Leu Asn Leu Lys Glu Ser Gly Val Asp Val Ile Val Gly Val Arg  
 35 40 45  
 Gln Gly Lys Ser Phe Thr Gln Ala Gln Glu Asp Gly His Lys Val Phe  
 50 55 60  
 Ser Val Lys Glu Ala Ala Ala Gln Ala Glu Ile Ile Met Val Leu Leu  
 65 70 75 80  
 Pro Asp Glu Gln Gln Gln Lys Val Tyr Glu Ala Glu Ile Lys Asp Glu  
 85 90 95  
 Leu Thr Ala Gly Lys Ser Leu Val Phe Ala His Gly Phe Asn Val His  
 100 105 110  
 Phe His Gln Ile Val Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala  
 115 120 125  
 Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Glu Gln Gly Ala  
 130 135 140  
 Gly Val Pro Ala Leu Phe Ala Ile Tyr Gln Asp Val Thr Gly Glu Ala  
 145 150 155 160  
 Arg Asp Lys Ala Leu Ala Tyr Ala Lys Gly Ile Gly Gly Ala Arg Ala  
 165 170 175  
 Gly Val Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe  
 180 185 190  
 Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Ser Ala Leu Val Lys Ala  
 195 200 205  
 Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr  
 210 215 220  
 Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu  
 225 230 235 240  
 Glu Gly Leu Ala Gly Met Arg Tyr Ser Ile Ser Asp Thr Ala Gln Trp  
 245 250 255  
 Gly Asp Phe Val Ser Gly Pro Arg Val Val Asp Ala Lys Val Lys Glu  
 260 265 270  
 Ser Met Lys Glu Val Leu Lys Asp Ile Gln Asn Gly Thr Phe Ala Lys  
 275 280 285  
 Glu Trp Ile Val Glu Asn Gln Val Asn Arg Pro Arg Phe Asn Ala Ile  
 290 295 300  
 Asn Ala Ser Glu Asn Glu His Gln Ile Glu Val Val Gly Arg Lys Leu  
 305 310 315 320  
 Arg Glu Met Met Pro Phe Val Lys Gln Gly Lys Lys Lys Glu Ala Val  
 325 330 335  
 Val Ser Val Ala Gln Asn

- 44 -

340

<210> 37  
 <211> 1674  
 <212> DNA  
 <213> Bacillus subtilis

<220>  
 <221> CDS  
 <222> (1)..(1674)

&lt;400&gt; 37

atg gca gaa tta cgc agt aat	atg atc aca caa gga atc gat aga gct	48
Met Ala Glu Leu Arg Ser Asn	Met Ile Thr Gln Gly Ile Asp Arg Ala	
1 5	10 15	
ccg cac cgc agt ttg ctt cgt	gca gca ggg gta aaa gaa gag gat ttc	96
Pro His Arg Ser Leu Leu Arg	Ala Ala Gly Val Lys Glu Glu Asp Phe	
20 25	30	
ggc aag ccg ttt att gcg gtg	tgt aat tca tac att gat atc gtt ccc	144
Gly Lys Pro Phe Ile Ala Val	Cys Asn Ser Tyr Ile Asp Ile Val Pro	
35 40	45	
ggt cat gtt cac ttg cag gag	ttt ggg aaa atc gta aaa gaa gca atc	192
Gly His Val His Leu Gln Glu	Phe Gly Lys Ile Val Lys Glu Ala Ile	
50 55	60	
aga gaa gca ggg ggc gtt ccg	ttt gaa ttt aat acc att ggg gta gat	240
Arg Glu Ala Gly Gly Val Pro	Phe Glu Phe Asn Thr Ile Gly Val Asp	
65 70	75 80	
gat ggc atc gca atg ggg cat	atc ggt atg aga tat tcg ctg cca agc	288
Asp Gly Ile Ala Met Gly His	Ile Gly Met Arg Tyr Ser Leu Pro Ser	
85 90	95	
cgt gaa att atc gca gac tct	gtg gaa acg gtt gta tcc gca cac tgg	336
Arg Glu Ile Ile Ala Asp Ser	Val Glu Thr Val Val Ser Ala His Trp	
100 105	110	
ttt gac gga atg gtc tgt att	ccg aac tgc gac aaa atc aca ccg gga	384
Phe Asp Gly Met Val Cys Ile	Pro Asn Cys Asp Lys Ile Thr Pro Gly	
115 120	125	
atg ctt atg gcg gca atg cgc	atc aac att ccg acg att ttt gtc agc	432
Met Leu Met Ala Ala Met Arg	Ile Asn Ile Pro Thr Ile Phe Val Ser	
130 135	140	
ggc gga ccg atg gcg gca gga	aga aca agt tac ggg cga aaa atc tcc	480
Gly Gly Pro Met Ala Ala Gly	Arg Thr Ser Tyr Gly Arg Lys Ile Ser	
145 150	155 160	
ctt tcc tca gta ttc gaa ggg	gta ggc gcc tac caa gca ggg aaa atc	528
Leu Ser Ser Val Phe Glu Gly	Val Gly Ala Tyr Gln Ala Gly Lys Ile	
165 170	175	
aac gaa aac gag ctt caa gaa	cta gag cag ttc gga tgc cca acg tgc	576

- 45 -

Asn	Glu	Asn	Glu	Leu	Gln	Glu	Leu	Glu	Gln	Phe	Gly	Cys	Pro	Thr	Cys	
		180						185					190			
ggg	tct	tgc	tca	ggc	atg	ttt	acg	gcg	aac	tca	atg	aac	tgt	ctg	tca	624
Gly	Ser	Cys	Ser	Gly	Met	Phe	Thr	Ala	Asn	Ser	Met	Asn	Cys	Leu	Ser	
		195					200					205				
gaa	gca	ctt	ggt	ctt	gct	ttg	ccg	ggt	aat	gga	acc	att	ctg	gca	aca	672
Glu	Ala	Leu	Gly	Leu	Ala	Leu	Pro	Gly	Asn	Gly	Thr	Ile	Leu	Ala	Thr	
	210					215					220					
tct	ccg	gaa	cgc	aaa	gag	ttt	gtg	aga	aaa	tcg	gct	gcg	caa	tta	atg	720
Ser	Pro	Glu	Arg	Lys	Glu	Phe	Val	Arg	Lys	Ser	Ala	Ala	Gln	Leu	Met	
225					230					235					240	
gaa	acg	att	cgc	aaa	gat	atc	aaa	ccg	cgt	gat	att	gtt	aca	gta	aaa	768
Glu	Thr	Ile	Arg	Lys	Asp	Ile	Lys	Pro	Arg	Asp	Ile	Val	Thr	Val	Lys	
			245						250				255			
gcg	att	gat	aac	gcg	ttt	gca	ctc	gat	atg	gcg	ctc	gga	ggt	tct	aca	816
Ala	Ile	Asp	Asn	Ala	Phe	Ala	Leu	Asp	Met	Ala	Leu	Gly	Gly	Ser	Thr	
			260					265					270			
aat	acc	gtt	ctt	cat	acc	ctt	gcc	ctt	gca	aac	gaa	gcc	ggc	gtt	gaa	864
Asn	Thr	Val	Leu	His	Thr	Leu	Ala	Leu	Ala	Asn	Glu	Ala	Gly	Val	Glu	
		275					280					285				
tac	tct	tta	gaa	cgc	att	aac	gaa	gtc	gct	gag	cgc	gtg	ccg	cac	ttg	912
Tyr	Ser	Leu	Glu	Arg	Ile	Asn	Glu	Val	Ala	Glu	Arg	Val	Pro	His	Leu	
	290					295					300					
gct	aag	ctg	gcg	cct	gca	tcg	gat	gtg	ttt	att	gaa	gat	ctt	cac	gaa	960
Ala	Lys	Leu	Ala	Pro	Ala	Ser	Asp	Val	Phe	Ile	Glu	Asp	Leu	His	Glu	
305					310					315					320	
gcg	ggc	ggc	gtt	tca	gcg	gct	ctg	aat	gag	ctt	tcg	aag	aaa	gaa	gga	1008
Ala	Gly	Gly	Val	Ser	Ala	Ala	Leu	Asn	Glu	Leu	Ser	Lys	Lys	Glu	Gly	
			325					330						335		
gcg	ctt	cat	tta	gat	gcg	ctg	act	gtt	aca	gga	aaa	act	ctt	gga	gaa	1056
Ala	Leu	His	Leu	Asp	Ala	Leu	Thr	Val	Thr	Gly	Lys	Thr	Leu	Gly	Glu	
			340				345						350			
acc	att	gcc	gga	cat	gaa	gta	aag	gat	tat	gac	gtc	att	cac	ccg	ctg	1104
Thr	Ile	Ala	Gly	His	Glu	Val	Lys	Asp	Tyr	Asp	Val	Ile	His	Pro	Leu	
		355					360					365				
gat	caa	cca	ttc	act	gaa	aag	gga	ggc	ctt	gct	gtt	tta	ttc	ggt	aat	1152
Asp	Gln	Pro	Phe	Thr	Glu	Lys	Gly	Gly	Leu	Ala	Val	Leu	Phe	Gly	Asn	
	370					375					380					
cta	gct	ccg	gac	ggc	gct	atc	att	aaa	aca	ggc	ggc	gta	cag	aat	ggg	1200
Leu	Ala	Pro	Asp	Gly	Ala	Ile	Ile	Lys	Thr	Gly	Gly	Val	Gln	Asn	Gly	
385					390					395					400	
att	aca	aga	cac	gaa	ggg	ccg	gct	gtc	gta	ttc	gat	tct	cag	gac	gag	1248
Ile	Thr	Arg	His	Glu	Gly	Pro	Ala	Val	Val	Phe	Asp	Ser	Gln	Asp	Glu	
			405					410						415		

- 46 -

gcg ctt gac ggc att atc aac cga aaa gta aaa gaa ggc gac gtt gtc 1296  
Ala Leu Asp Gly Ile Ile Asn Arg Lys Val Lys Glu Gly Asp Val Val  
420 425 430

atc atc aga tac gaa ggg cca aaa ggc gga cct ggc atg ccg gaa atg 1344  
Ile Ile Arg Tyr Glu Gly Pro Lys Gly Gly Pro Gly Met Pro Glu Met  
435 440 445

ctg gcg cca aca tcc caa atc gtt gga atg gga ctc ggc cca aaa gtg 1392  
Leu Ala Pro Thr Ser Gln Ile Val Gly Met Gly Leu Gly Pro Lys Val  
450 455 460

gca ttg att acg gac gga cgt ttt tcc gga gcc tcc cgt ggc ctc tca 1440  
Ala Leu Ile Thr Asp Gly Arg Phe Ser Gly Ala Ser Arg Gly Leu Ser  
465 470 475 480

atc ggc cac gta tca cct gag gcc gct gag ggc ggg ccg ctt gcc ttt 1488  
Ile Gly His Val Ser Pro Glu Ala Ala Glu Gly Gly Pro Leu Ala Phe  
485 490 495

gtt gaa aac gga gac cat att atc gtt gat att gaa aaa cgc atc ttg 1536  
Val Glu Asn Gly Asp His Ile Ile Val Asp Ile Glu Lys Arg Ile Leu  
500 505 510

gat gta caa gtg cca gaa gaa gag tgg gaa aaa cga aaa gcg aac tgg 1584  
Asp Val Gln Val Pro Glu Glu Glu Trp Glu Lys Arg Lys Ala Asn Trp  
515 520 525

aaa ggt ttt gaa ccg aaa gtg aaa acc ggc tac ctg gca cgt tat tct 1632  
Lys Gly Phe Glu Pro Lys Val Lys Thr Gly Tyr Leu Ala Arg Tyr Ser  
530 535 540

aaa ctt gtg aca agt gcc aac acc ggc ggt att atg aaa atc 1674  
Lys Leu Val Thr Ser Ala Asn Thr Gly Gly Ile Met Lys Ile  
545 550 555

&lt;210&gt; 38

&lt;211&gt; 558

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 38

Met Ala Glu Leu Arg Ser Asn Met Ile Thr Gln Gly Ile Asp Arg Ala  
1 5 10 15

Pro His Arg Ser Leu Leu Arg Ala Ala Gly Val Lys Glu Glu Asp Phe  
20 25 30

Gly Lys Pro Phe Ile Ala Val Cys Asn Ser Tyr Ile Asp Ile Val Pro  
35 40 45

Gly His Val His Leu Gln Glu Phe Gly Lys Ile Val Lys Glu Ala Ile  
50 55 60

Arg Glu Ala Gly Gly Val Pro Phe Glu Phe Asn Thr Ile Gly Val Asp  
65 70 75 80

- 47 -

Asp	Gly	Ile	Ala	Met	Gly	His	Ile	Gly	Met	Arg	Tyr	Ser	Leu	Pro	Ser		
				85					90					95			
Arg	Glu	Ile	Ile	Ala	Asp	Ser	Val	Glu	Thr	Val	Val	Ser	Ala	His	Trp		
			100					105					110				
Phe	Asp	Gly	Met	Val	Cys	Ile	Pro	Asn	Cys	Asp	Lys	Ile	Thr	Pro	Gly		
		115					120					125					
Met	Leu	Met	Ala	Ala	Met	Arg	Ile	Asn	Ile	Pro	Thr	Ile	Phe	Val	Ser		
	130					135					140						
Gly	Gly	Pro	Met	Ala	Ala	Gly	Arg	Thr	Ser	Tyr	Gly	Arg	Lys	Ile	Ser		
145					150					155					160		
Leu	Ser	Ser	Val	Phe	Glu	Gly	Val	Gly	Ala	Tyr	Gln	Ala	Gly	Lys	Ile		
				165					170					175			
Asn	Glu	Asn	Glu	Leu	Gln	Glu	Leu	Glu	Gln	Phe	Gly	Cys	Pro	Thr	Cys		
			180					185					190				
Gly	Ser	Cys	Ser	Gly	Met	Phe	Thr	Ala	Asn	Ser	Met	Asn	Cys	Leu	Ser		
		195					200					205					
Glu	Ala	Leu	Gly	Leu	Ala	Leu	Pro	Gly	Asn	Gly	Thr	Ile	Leu	Ala	Thr		
	210					215					220						
Ser	Pro	Glu	Arg	Lys	Glu	Phe	Val	Arg	Lys	Ser	Ala	Ala	Gln	Leu	Met		
225					230					235					240		
Glu	Thr	Ile	Arg	Lys	Asp	Ile	Lys	Pro	Arg	Asp	Ile	Val	Thr	Val	Lys		
				245					250					255			
Ala	Ile	Asp	Asn	Ala	Phe	Ala	Leu	Asp	Met	Ala	Leu	Gly	Gly	Ser	Thr		
			260					265					270				
Asn	Thr	Val	Leu	His	Thr	Leu	Ala	Leu	Ala	Asn	Glu	Ala	Gly	Val	Glu		
		275					280					285					
Tyr	Ser	Leu	Glu	Arg	Ile	Asn	Glu	Val	Ala	Glu	Arg	Val	Pro	His	Leu		
	290					295					300						
Ala	Lys	Leu	Ala	Pro	Ala	Ser	Asp	Val	Phe	Ile	Glu	Asp	Leu	His	Glu		
305					310					315					320		
Ala	Gly	Gly	Val	Ser	Ala	Ala	Leu	Asn	Glu	Leu	Ser	Lys	Lys	Glu	Gly		
				325				330						335			
Ala	Leu	His	Leu	Asp	Ala	Leu	Thr	Val	Thr	Gly	Lys	Thr	Leu	Gly	Glu		
			340					345					350				
Thr	Ile	Ala	Gly	His	Glu	Val	Lys	Asp	Tyr	Asp	Val	Ile	His	Pro	Leu		
		355					360					365					
Asp	Gln	Pro	Phe	Thr	Glu	Lys	Gly	Gly	Leu	Ala	Val	Leu	Phe	Gly	Asn		
	370					375					380						

- 48 -

Leu Ala Pro Asp Gly Ala Ile Ile Lys Thr Gly Gly Val Gln Asn Gly  
 385 390 395 400  
 Ile Thr Arg His Glu Gly Pro Ala Val Val Phe Asp Ser Gln Asp Glu  
 405 410 415  
 Ala Leu Asp Gly Ile Ile Asn Arg Lys Val Lys Glu Gly Asp Val Val  
 420 425 430  
 Ile Ile Arg Tyr Glu Gly Pro Lys Gly Gly Pro Gly Met Pro Glu Met  
 435 440 445  
 Leu Ala Pro Thr Ser Gln Ile Val Gly Met Gly Leu Gly Pro Lys Val  
 450 455 460  
 Ala Leu Ile Thr Asp Gly Arg Phe Ser Gly Ala Ser Arg Gly Leu Ser  
 465 470 475 480  
 Ile Gly His Val Ser Pro Glu Ala Ala Glu Gly Gly Pro Leu Ala Phe  
 485 490 495  
 Val Glu Asn Gly Asp His Ile Ile Val Asp Ile Glu Lys Arg Ile Leu  
 500 505 510  
 Asp Val Gln Val Pro Glu Glu Glu Trp Glu Lys Arg Lys Ala Asn Trp  
 515 520 525  
 Lys Gly Phe Glu Pro Lys Val Lys Thr Gly Tyr Leu Ala Arg Tyr Ser  
 530 535 540  
 Lys Leu Val Thr Ser Ala Asn Thr Gly Gly Ile Met Lys Ile  
 545 550 555

<210> 39  
 <211> 194  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: promoter  
 sequence

<220>  
 <221> -35\_signal  
 <222> (136)..(141)

<220>  
 <221> -10\_signal  
 <222> (159)..(164)

<400> 39  
 gctattgacg acagctatgg ttcaactgtcc accaaccaaa actgtgctca gtaccgccaa 60  
 tattttctccc ttgaggggta caaagaggtg tccctagaag agatccacgc tgtgtaaaaa 120  
 ttttacaaaa aggtattgac tttccctaca ggggtgtgtaa taatttaatt acaggcgggg 180

- 49 -

gcaaccccgctgt

194

&lt;210&gt; 40

&lt;211&gt; 163

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:promoter  
sequence

&lt;220&gt;

&lt;221&gt; -35\_signal

&lt;222&gt; (113)..(118)

&lt;220&gt;

&lt;221&gt; -10\_signal

&lt;222&gt; (136)..(141)

&lt;400&gt; 40

gcctacctag cttccaagaa agatattcta acagcacaag agcggaaaga tgttttggtc 60

tacatccaga acaacctctg ctaaaattcc tgaaaaattt tgcaaaaagt tgttgacttt 120

atctacaagg tgtggtataa taatcttaac aacagcagga cgc

163

&lt;210&gt; 41

&lt;211&gt; 127

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:promoter  
sequence

&lt;220&gt;

&lt;221&gt; -35\_signal

&lt;222&gt; (34)..(39)

&lt;220&gt;

&lt;221&gt; -10\_signal

&lt;222&gt; (58)..(63)

&lt;220&gt;

&lt;221&gt; -35\_signal

&lt;222&gt; (75)..(80)

&lt;220&gt;

&lt;221&gt; -10\_signal

&lt;222&gt; (98)..(103)

&lt;400&gt; 41

gaggaatcat agaattttgt caaaataatt ttattgacaa cgtcttatta acgttgatat 60

aatttaaatt ttatttgaca aaaatgggct cgtgttgtag aataaatgta gtgaggtgga 120



tgcaatg

127

&lt;210&gt; 42

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 42

taaacatgag gaggagaaaa catg

24

&lt;210&gt; 43

&lt;211&gt; 28

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 43

attcgagaaa tggagagaat ataatatg

28

&lt;210&gt; 44

&lt;211&gt; 13

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 44

agaaaggagg tga

13

&lt;210&gt; 45

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 45

ttaagaaagg aggtgannnn atg

23

&lt;210&gt; 46

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence:ribosome

- 51 -

## binding site

<400> 46  
ttagaaagga ggtgannnnn atg 23

<210> 47  
<211> 23  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence:ribosome  
binding site

<400> 47  
agaaaggagg tgannnnnnn atg 23

<210> 48  
<211> 22  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence:ribosome  
binding site

<400> 48  
agaaaggagg tgannnnnna tg 22

<210> 49  
<211> 25  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence:ribosome  
binding site

<400> 49  
ccctctagaa ggaggagaaa acatg 25

<210> 50  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence:ribosome  
binding site

<400> 50  
ccctctagag gaggagaaaa catg 24

<210> 51  
<211> 23  
<212> DNA

- 52 -

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 51

ttagaaagga ggattttaa atg

23

&lt;210&gt; 52

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 52

ttagaaagga gggttaatta atg

23

&lt;210&gt; 53

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 53

ttagaaagga ggtgatttaa atg

23

&lt;210&gt; 54

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 54

ttagaaagga ggtgtttaa atg

23

&lt;210&gt; 55

&lt;211&gt; 28

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 55

attcgagaaa ggaggtgaat ataatatg

28

&lt;210&gt; 56

- 53 -

&lt;211&gt; 27

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 56

attcgagaaa ggaggtgaat aataatg

27

&lt;210&gt; 57

&lt;211&gt; 28

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:ribosome  
binding site

&lt;400&gt; 57

attcgtagaa aggagtgaa ttaatatg

28

&lt;210&gt; 58

&lt;211&gt; 3291

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 58

atggggacta atgtacaggt ggattcagca tctgccgaat gtacacagac gatgagcgga 60

gcattaatgc tgattgaatc attaaaaaaa gagaaagtag aaatgatctt cggttatccg 120

ggcggggctg tgcttccgat ttacgataag ctatacaatt cagggttggt acatatacctt 180

ccccgtcacg aacaaggagc aattcatgca gcggagggat acgcaagggt ctccggaaaa 240

ccgggtgtcg tcattgccac gtcagggccg ggagcgacaa accttggttac aggccttgct 300

gatgccatga ttgattcatt gccgtagtc gtctttacag ggcaggtagc aacctctgta 360

atcgggagcg atgcatttca ggaagcagac attttaggga ttacgatgcc agtaacaaaa 420

cacagctacc aggttcgcca gccggaagat ctgccgcgca tcattaaaga agcgttccat 480

attgcaacaa ctggaagacc cggacctgta ttgattgata ttccgaaaga ttagtagcaaca 540

attgaaggag aattcagcta cgatcatgag atgaatctcc cgggatacca gccgacaaca 600

gagccgaatt atttcagat ccgcaagctt gtggaagccg tgagcagtgc gaaaaaacg 660

gtgatcctgg cgggtgcggg cgtactgcac ggaaaagcgt cagaagaatt aaaaaattat 720

gctgaacagc agcaaattccc tgtggcacac acccttttgg ggctcggagg cttcccggct 780

gaccatccgc ttttcctagg gatggcgga atgcacgga cttatacagc caatatggcc 840

cttcatgaat gtgatctatt aatcagtatc ggcgcccggt ttgatgaccg tgtcacagga 900  
aacctgaaac actttgccag aaacgcaaag atagcccaca tcgatattga tccagctgaa 960  
atcggaaaaa tcatgaaaac acagattcct gtagtcggag acagcaaaat tgtcctgcag 1020  
gagctgatca aacaagacgg caaacaagc gattcaagcg aatggaaaaa acagctcgca 1080  
gaatggaaag aagagtatcc gctctggat gtagataatg aagaagaagg ttttaaacct 1140  
cagaaattga ttgaatatat tcatcaattt acaaaaggag aggccattgt cgcaacggat 1200  
gtaggccagc atcaaatgtg gtcagcgcaa ttttatccgt tccaaaaagc agataaatgg 1260  
gtcacgtcag gcggaacttg aacgatggga ttccggtcttc cggcggcgat cggcgcacag 1320  
ctggccgaaa aagatgctac tgttgctcgcg gttgtcggag acggcggatt ccaaagacg 1380  
cttcaagaac tcgatgttat tcgcgaatta aatcttccgg tcaaggtagt gattttaaat 1440  
aacgcttgct tcggaatggt cagacagtgg caggaaattt tctatgaaga acgttattca 1500  
gaatctaaat tcgcttctca gcctgacttc gtcaaattgt ccgaagcata cggcattaaa 1560  
ggcatcagaa tttcatcaga agcgggaagc aaggaaaagc tggaagaggc attaacatca 1620  
agagaacctg ttgtcattga cgtgcgggtt gccagcgaag aaaaagtatt cccgatggtg 1680  
gctccgggga aagggtgca tgaaatggtg ggggtgaaac cttgaaaaga attatcacat 1740  
tgactgtggt gaaccgctcc ggggtgttaa accggatcac cgggtctattc acaaaaaggc 1800  
attacaacat tgaaagcatt acagttggac acacagaaac agccggcggt tccagaatca 1860  
ccttcgtcgt tcatgttgaa ggtgaaaatg atgttgaaca gttaacgaaa cagctcaaca 1920  
aacagattga tgtgctgaaa gtcacagaca tcacaaatca atcgattgtc cagagggagc 1980  
tggccttaat caaggttgct tccgcacctt caacaagaac agagattaat ggaatcatag 2040  
aaccgtttag agcctctgtc gttgatgtca gcagagacag catcgttgtt caggtgacag 2100  
gtgaatctaa caaaattgaa gcgcttattg agttattaaa accttatggc attaaagaaa 2160  
tcgcgagaac aggtacaacg gcttttgca ggggaaccag caaaaggcgt catccaataa 2220  
aacaatatct attgtataaa acataacaag ggagagattg aaatggtaaa agtatattat 2280  
aacggtgata tcaaagagaa cgtattggct ggaaaaacag tagcggttat cgggtacggg 2340  
tcgcaaggcc acgcacatgc cctgaacctt aaagaaagcg gagtagacgt gatcgctcgg 2400  
gtagacaag gaaaatcttt cactcaagcc caagaagacg gacataaagt attttcagta 2460  
aaagaagcgg cagcccaagc cgaaatcatc atggttctgc ttccggatga gcagcagcaa 2520  
aaagtatacg aagctgaaat caaagatgaa ttgacagcag gaaaatcatt agtatcgcgt 2580

- 55 -

catggattta acgtgcattt ccatcaaatt gttcctccgg cggatgtaga tgtattctta 2640  
 gtggccccta aaggcccggg acacttggtg agaagaacat atgagcaagg agctggcgta 2700  
 cctgcattgt tgcgaatcta tcaagatgtg actggagaag caagagacaa agccctcgct 2760  
 tatgctaaag gaatcggcgg cgcaagagcg ggcgtattag aaacgacatt taaagaagaa 2820  
 acagaaacag atttgttcgg tgagcaagca gttctttgcg gcggattaag cgcgcttgtc 2880  
 aaagccggat ttgaaacctt aactgaagca ggttatcagc ctgaacttgc atacttcgag 2940  
 tgtcttcattg agctgaaatt aatcgtagac cttatgtacg aagaaggact tgcaggaatg 3000  
 agatattcaa tctctgacac agcacagtgg ggagatttcg tatcaggccc tcgcgttggtg 3060  
 gacgccaaag taaaagaatc tatgaaagaa gtattaaaag atatccaaaa cggtagattc 3120  
 gcaaaagagt ggatcgtcga aaaccaagta aaccgtcctc gtttcaacgc tatcaatgca 3180  
 agcgagaacg aacatcaaatt cgaagtagtg ggaagaaagc ttcgtgaaat gatgccgttt 3240  
 gtgaaacaag gcaagaagaa ggaagcggtg gtctccgttg cgcaaaatta a 3291

<210> 59  
 <211> 2363  
 <212> DNA  
 <213> Bacillus subtilis

<220>  
 <221> CDS  
 <222> (242)..(1072)

<220>  
 <221> CDS  
 <222> (1077)..(1934)

<220>  
 <221> CDS  
 <222> (1939)..(2319)

<400> 59  
 ttggtacaag cccgttgatt ttggtatact tccattgggc agtatgcct gcgaactgca 60  
 cctattatta aaatagatag acattgcagc agtctgcctt gatccaaaaa aggactggga 120  
 cagagggatg aaactcgccg aactttagaa agtgaagaat ccttctcggt gtaacggaag 180  
 gttttttggc ttgcagaaga aaacggcaga tcatctcctc taaacatgag gaggagaaaa 240  
 c atg aaa aca aaa ctg gat ttt cta aaa atg aag gag tct gaa gaa ccg 289  
 Met Lys Thr Lys Leu Asp Phe Leu Lys Met Lys Glu Ser Glu Glu Pro  
 1 5 10 15

att gtc atg ctg acc gct tat gat tat ccg gca gct aaa ctt gct gaa 337  
 Ile Val Met Leu Thr Ala Tyr Asp Tyr Pro Ala Ala Lys Leu Ala Glu

- 56 -

20								25								30								
caa	gcg	gga	gtt	gac	atg	att	tta	gtc	ggg	gat	tca	ctt	gga	atg	gtc									385
Gln	Ala	Gly	Val	Asp	Met	Ile	Leu	Val	Gly	Asp	Ser	Leu	Gly	Met	Val									
		35					40					45												
gtc	ctc	ggc	ctt	gat	tca	act	gtc	ggg	gtg	aca	gtt	gcg	gac	atg	atc									433
Val	Leu	Gly	Leu	Asp	Ser	Thr	Val	Gly	Val	Thr	Val	Ala	Asp	Met	Ile									
		50				55					60													
cat	cat	aca	aaa	gcc	gtt	aaa	agg	ggg	gcg	ccg	aat	acc	ttt	att	gtg									481
His	His	Thr	Lys	Ala	Val	Lys	Arg	Gly	Ala	Pro	Asn	Thr	Phe	Ile	Val									
		65				70				75					80									
aca	gat	atg	ccg	ttt	atg	tct	tat	cac	ctg	tct	aag	gaa	gat	acg	ctg									529
Thr	Asp	Met	Pro	Phe	Met	Ser	Tyr	His	Leu	Ser	Lys	Glu	Asp	Thr	Leu									
				85					90					95										
aaa	aat	gca	gcg	gct	atc	gtt	cag	gaa	agc	gga	gct	gac	gca	ctg	aag									577
Lys	Asn	Ala	Ala	Ala	Ile	Val	Gln	Glu	Ser	Gly	Ala	Asp	Ala	Leu	Lys									
			100				105					110												
ctt	gag	ggc	gga	gaa	ggc	gtg	ttt	gaa	tcc	att	cgc	gca	ttg	acg	ctt									625
Leu	Glu	Gly	Gly	Glu	Gly	Val	Phe	Glu	Ser	Ile	Arg	Ala	Leu	Thr	Leu									
		115					120					125												
gga	ggc	att	cca	gta	gtc	agt	cac	tta	ggg	ttg	aca	ccg	cag	tca	gtc									673
Gly	Gly	Ile	Pro	Val	Val	Ser	His	Leu	Gly	Leu	Thr	Pro	Gln	Ser	Val									
		130				135					140													
ggc	gta	ctg	ggc	ggc	tat	aaa	gta	cag	ggc	aaa	gac	gaa	caa	agc	gcc									721
Gly	Val	Leu	Gly	Gly	Tyr	Lys	Val	Gln	Gly	Lys	Asp	Glu	Gln	Ser	Ala									
		145			150				155						160									
aaa	aaa	tta	ata	gaa	gac	agt	ata	aaa	tgc	gaa	gaa	gca	gga	gct	atg									769
Lys	Lys	Leu	Ile	Glu	Asp	Ser	Ile	Lys	Cys	Glu	Glu	Ala	Gly	Ala	Met									
				165					170					175										
atg	ctt	gtg	ctg	gaa	tgt	gtg	ccg	gca	gaa	ctc	aca	gcc	aaa	att	gcc									817
Met	Leu	Val	Leu	Glu	Cys	Val	Pro	Ala	Glu	Leu	Thr	Ala	Lys	Ile	Ala									
			180					185					190											
gag	acg	cta	agc	ata	ccg	gtc	att	gga	atc	ggg	gct	ggg	gtg	aaa	gcg									865
Glu	Thr	Leu	Ser	Ile	Pro	Val	Ile	Gly	Ile	Gly	Ala	Gly	Val	Lys	Ala									
		195					200					205												
gac	gga	caa	gtt	ctc	gtt	tat	cat	gat	att	atc	ggc	cac	ggg	gtt	gag									913
Asp	Gly	Gln	Val	Leu	Val	Tyr	His	Asp	Ile	Ile	Gly	His	Gly	Val	Glu									
		210				215					220													
aga	aca	cct	aaa	ttt	gta	aag	caa	tat	acg	cgc	att	gat	gaa	acc	atc									961
Arg	Thr	Pro	Lys	Phe	Val	Lys	Gln	Tyr	Thr	Arg	Ile	Asp	Glu	Thr	Ile									
		225			230				235					240										
gaa	aca	gca	atc	agc	gga	tat	gtt	cag	gat	gta	aga	cat	cgt	gct	ttc									1009
Glu	Thr	Ala	Ile	Ser	Gly	Tyr	Val	Gln	Asp	Val	Arg	His	Arg	Ala	Phe									
				245				250						255										

- 57 -

cct gaa caa aag cat tcc ttt caa atg aac cag aca gtg ctt gac ggc	1057
Pro Glu Gln Lys His Ser Phe Gln Met Asn Gln Thr Val Leu Asp Gly	
260 265 270	
ttg tac ggg gga aaa taag atg aga cag att act gat att tca cag ctg	1106
Leu Tyr Gly Gly Lys Met Arg Gln Ile Thr Asp Ile Ser Gln Leu	
275 280 285	
aaa gaa gcc ata aaa caa tac cat tca gag ggc aag tca atc gga ttt	1154
Lys Glu Ala Ile Lys Gln Tyr His Ser Glu Gly Lys Ser Ile Gly Phe	
290 295 300	
gtt ccg acg atg ggg ttt ctg cat gag ggg cat tta acc tta gca gac	1202
Val Pro Thr Met Gly Phe Leu His Glu Gly His Leu Thr Leu Ala Asp	
305 310 315	
aaa gca aga caa gaa aac gac gcc gtt att atg agt att ttt gtg aat	1250
Lys Ala Arg Gln Glu Asn Asp Ala Val Ile Met Ser Ile Phe Val Asn	
320 325 330 335	
cct gca caa ttc ggc cct aat gaa gat ttt gaa gca tat ccg cgc gat	1298
Pro Ala Gln Phe Gly Pro Asn Glu Asp Phe Glu Ala Tyr Pro Arg Asp	
340 345 350	
att gag cgg gat gca gct ctt gca gaa aac gcc gga gtc gat att ctt	1346
Ile Glu Arg Asp Ala Ala Leu Ala Glu Asn Ala Gly Val Asp Ile Leu	
355 360 365	
ttt acg cca gat gct cat gat atg tat ccc ggt gaa aag aat gtc acg	1394
Phe Thr Pro Asp Ala His Asp Met Tyr Pro Gly Glu Lys Asn Val Thr	
370 375 380	
att cat gta gaa aga cgc aca gac gtg tta tgc ggg cgc tca aga gaa	1442
Ile His Val Glu Arg Arg Thr Asp Val Leu Cys Gly Arg Ser Arg Glu	
385 390 395	
gga cat ttt gac ggg gtc gcg atc gta ctg acg aag ctt ttc aat cta	1490
Gly His Phe Asp Gly Val Ala Ile Val Leu Thr Lys Leu Phe Asn Leu	
400 405 410 415	
gtc aag ccg act cgt gcc tat ttc ggt tta aaa gat gcg cag cag gta	1538
Val Lys Pro Thr Arg Ala Tyr Phe Gly Leu Lys Asp Ala Gln Gln Val	
420 425 430	
gct gtt gtt gat ggg tta atc agc gac ttc ttc atg gat att gaa ttg	1586
Ala Val Val Asp Gly Leu Ile Ser Asp Phe Phe Met Asp Ile Glu Leu	
435 440 445	
gtt cct gtc gat acg gtc aga gag gaa gac ggc tta gcc aaa agc tct	1634
Val Pro Val Asp Thr Val Arg Glu Glu Asp Gly Leu Ala Lys Ser Ser	
450 455 460	
cgc aat gta tac tta aca gct gag gaa aga aaa gaa gcg cct aag ctg	1682
Arg Asn Val Tyr Leu Thr Ala Glu Glu Arg Lys Glu Ala Pro Lys Leu	
465 470 475	
tat ccg gcc ctt caa aca agt gcg gaa ctt gtc caa gcc ggt gaa aga	1730
Tyr Arg Ala Leu Gln Thr Ser Ala Glu Leu Val Gln Ala Gly Glu Arg	



- 58 -

480	485	490	495	
gat cct gaa gcg gtg ata aaa gct gca aaa gat atc att gaa acg act				1778
Asp Pro Glu Ala Val Ile Lys Ala Ala Lys Asp Ile Ile Glu Thr Thr				
	500	505	510	
agc gga acc ata gac tat gta gag ctt tat tcc tat ccg gaa ctc gag				1826
Ser Gly Thr Ile Asp Tyr Val Glu Leu Tyr Ser Tyr Pro Glu Leu Glu				
	515	520	525	
cct gtg aat gaa att gct gga aag atg att ctc gct gtt gca gtt gct				1874
Pro Val Asn Glu Ile Ala Gly Lys Met Ile Leu Ala Val Ala Val Ala				
	530	535	540	
ttt tca aaa gcg cgt tta ata gat aat atc att att gat att cga gaa				1922
Phe Ser Lys Ala Arg Leu Ile Asp Asn Ile Ile Ile Asp Ile Arg Glu				
	545	550	555	
atg gag aga ata taat atg tat cga aca atg atg agc ggc aaa ctt cac				1971
Met Glu Arg Ile Met Tyr Arg Thr Met Met Ser Gly Lys Leu His				
	560	565	570	
agg gca act gtt acg gaa gca aac ctg aac tat gtg gga agc att aca				2019
Arg Ala Thr Val Thr Glu Ala Asn Leu Asn Tyr Val Gly Ser Ile Thr				
	575	580	585	590
att gat gaa gat ctc att gat gct gtg gga atg ctt cct aat gaa aaa				2067
Ile Asp Glu Asp Leu Ile Asp Ala Val Gly Met Leu Pro Asn Glu Lys				
	595	600	605	
gta caa att gtg aat aat aat aat gga gca cgt ctt gaa acg tat att				2115
Val Gln Ile Val Asn Asn Asn Asn Gly Ala Arg Leu Glu Thr Tyr Ile				
	610	615	620	
att cct ggt aaa cgg gga agc ggc gtc ata tgc tta aac ggt gca gcc				2163
Ile Pro Gly Lys Arg Gly Ser Gly Val Ile Cys Leu Asn Gly Ala Ala				
	625	630	635	
gca cgc ctt gtg cag gaa gga gat aag gtc att att att tcc tac aaa				2211
Ala Arg Leu Val Gln Glu Gly Asp Lys Val Ile Ile Ile Ser Tyr Lys				
	640	645	650	
atg atg tct gat caa gaa gcg gca agc cat gag ccg aaa gtg gct gtt				2259
Met Met Ser Asp Gln Glu Ala Ala Ser His Glu Pro Lys Val Ala Val				
	655	660	665	670
ctg aat gat caa aac aaa att gaa caa atg ctg ggg aac gaa cca gcc				2307
Leu Asn Asp Gln Asn Lys Ile Glu Gln Met Leu Gly Asn Glu Pro Ala				
	675	680	685	
cgt aca att ttg tagaagaaaa gcccccttta tcgggggttt tcttttaaga tttt				2363
Arg Thr Ile Leu				
	690			

<210> 60  
 <211> 293  
 <212> PRT

- 59 -

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 60

```

Met Ser Ile Ala Val Ser Glu Glu Glu Ala Lys Ala Val Glu Gly Leu
 1           5           10           15

Asn Asp Tyr Leu Ser Val Glu Glu Val Glu Thr Ile Tyr Ile Pro Leu
          20           25           30

Val Arg Leu Leu His Leu His Val Lys Ser Ala Ala Glu Arg Asn Lys
          35           40           45

His Val Asn Val Phe Leu Lys His Pro His Ser Ala Lys Ile Pro Phe
 50           55           60

Ile Ile Gly Ile Ala Gly Ser Val Ala Val Gly Lys Ser Thr Thr Ala
65           70           75           80

Arg Ile Leu Gln Lys Leu Leu Ser Arg Leu Pro Asp Arg Pro Lys Val
          85           90           95

Ser Leu Ile Thr Thr Asp Gly Phe Leu Phe Pro Thr Ala Glu Leu Lys
          100          105          110

Lys Lys Asn Met Met Ser Arg Lys Gly Phe Pro Glu Ser Tyr Asp Val
          115          120          125

Lys Ala Leu Leu Glu Phe Leu Asn Asp Leu Lys Ser Gly Lys Asp Ser
          130          135          140

Val Lys Ala Pro Val Tyr Ser His Leu Thr Tyr Asp Arg Glu Glu Gly
          145          150          155          160

Val Phe Glu Val Val Glu Gln Ala Asp Ile Val Ile Ile Glu Gly Ile
          165          170          175

Asn Val Leu Gln Ser Pro Thr Leu Glu Asp Asp Arg Glu Asn Pro Arg
          180          185          190

Ile Phe Val Ser Asp Phe Phe Asp Phe Ser Ile Tyr Val Asp Ala Glu
          195          200          205

Glu Ser Arg Ile Phe Thr Trp Tyr Leu Glu Arg Phe Arg Leu Leu Arg
          210          215          220

Glu Thr Ala Phe Gln Asn Pro Asp Ser Tyr Phe His Lys Phe Lys Asp
          225          230          235          240

Leu Ser Asp Gln Glu Ala Asp Glu Met Ala Ala Ser Ile Trp Glu Ser
          245          250          255

Val Asn Arg Pro Asn Leu Tyr Glu Asn Ile Leu Pro Thr Lys Phe Arg
          260          265          270

Ser Asp Leu Ile Leu Arg Lys Gly Asp Gly His Lys Val Glu Glu Val
          275          280          285

Leu Val Arg Arg Val

```

- 60 -

290

&lt;210&gt; 61

&lt;211&gt; 281

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 61

Met	Glu	Gly	Leu	Asn	Asp	Tyr	Leu	Ser	Val	Glu	Glu	Val	Glu	Thr	Ile
1				5					10					15	

Tyr	Ile	Pro	Leu	Val	Arg	Leu	Leu	His	Leu	His	Val	Lys	Ser	Ala	Ala
			20					25					30		

Glu	Arg	Asn	Lys	His	Val	Asn	Val	Phe	Leu	Lys	His	Pro	His	Ser	Ala
		35					40						45		

Lys	Ile	Pro	Phe	Ile	Ile	Gly	Ile	Ala	Gly	Ser	Val	Ala	Val	Gly	Lys
50						55					60				

Ser	Thr	Thr	Ala	Arg	Ile	Leu	Gln	Lys	Leu	Leu	Ser	Arg	Leu	Pro	Asp
65					70					75					80

Arg	Pro	Lys	Val	Ser	Leu	Ile	Thr	Thr	Asp	Gly	Phe	Leu	Phe	Pro	Thr
				85					90					95	

Ala	Glu	Leu	Lys	Lys	Lys	Asn	Met	Met	Ser	Arg	Lys	Gly	Phe	Pro	Glu
			100					105					110		

Ser	Tyr	Asp	Val	Lys	Ala	Leu	Leu	Glu	Phe	Leu	Asn	Asp	Leu	Lys	Ser
		115					120					125			

Gly	Lys	Asp	Ser	Val	Lys	Ala	Pro	Val	Tyr	Ser	His	Leu	Thr	Tyr	Asp
130						135					140				

Arg	Glu	Glu	Gly	Val	Phe	Glu	Val	Val	Glu	Gln	Ala	Asp	Ile	Val	Ile
145					150					155					160

Ile	Glu	Gly	Ile	Asn	Val	Leu	Gln	Ser	Pro	Thr	Leu	Glu	Asp	Asp	Arg
				165					170					175	

Glu	Asn	Pro	Arg	Ile	Phe	Val	Ser	Asp	Phe	Phe	Asp	Phe	Ser	Ile	Tyr
			180					185					190		

Val	Asp	Ala	Glu	Glu	Ser	Arg	Ile	Phe	Thr	Trp	Tyr	Leu	Glu	Arg	Phe
		195					200					205			

Arg	Leu	Leu	Arg	Glu	Thr	Ala	Phe	Gln	Asn	Pro	Asp	Ser	Tyr	Phe	His
210						215					220				

Lys	Phe	Lys	Asp	Leu	Ser	Asp	Gln	Glu	Ala	Asp	Glu	Met	Ala	Ala	Ser
225					230					235					240

Ile	Trp	Glu	Ser	Val	Asn	Arg	Pro	Asn	Leu	Tyr	Glu	Asn	Ile	Leu	Pro
				245					250					255	

Thr	Lys	Phe	Arg	Ser	Asp	Leu	Ile	Leu	Arg	Lys	Gly	Asp	Gly	His	Lys
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- 61 -

260 265 270  
 Val Glu Glu Val Leu Val Arg Arg Val  
 275 280

<210> 62  
 <211> 1092  
 <212> DNA  
 <213> *Bacillus subtilis*  
 <220>  
 <221> CDS  
 <222> (1)..(1089)

<400> 62  
 atg act aaa caa acá att cgc gtt gaa ttg aca tca aca aaa aaa ccg 48  
 Met Thr Lys Gln Thr Ile Arg Val Glu Leu Thr Ser Thr Lys Lys Pro  
 1 5 10 15  
 aaa cca gac cca aat cag ctt tcg ttc gga aga gtg ttt aca gac cac 96  
 Lys Pro Asp Pro Asn Gln Leu Ser Phe Gly Arg Val Phe Thr Asp His  
 20 25 30  
 atg ttt gta atg gac tat gcc gca gat aaa ggt tgg tac gat cca aga 144  
 Met Phe Val Met Asp Tyr Ala Ala Asp Lys Gly Trp Tyr Asp Pro Arg  
 35 40 45  
 atc att cct tat caa ccc tta tca atg gat cca act gca atg gtc tat 192  
 Ile Ile Pro Tyr Gln Pro Leu Ser Met Asp Pro Thr Ala Met Val Tyr  
 50 55 60  
 cac tac ggc caa acc gtg ttt gaa ggg tta aag gct tac gtg tca gag 240  
 His Tyr Gly Gln Thr Val Phe Glu Gly Leu Lys Ala Tyr Val Ser Glu  
 65 70 75 80  
 gat gac cat gtt ctg ctt ttc aga ccg gaa aaa aat atg gaa cgc ctg 288  
 Asp Asp His Val Leu Leu Phe Arg Pro Glu Lys Asn Met Glu Arg Leu  
 85 90 95  
 aat caa tca aac gac cgc ctc tgc atc ccg caa att gat gaa gaa cag 336  
 Asn Gln Ser Asn Asp Arg Leu Cys Ile Pro Gln Ile Asp Glu Glu Gln  
 100 105 110  
 gtt ctt gaa ggc tta aag cag ctt gtc gca att gat aaa gac tgg att 384  
 Val Leu Glu Gly Leu Lys Gln Leu Val Ala Ile Asp Lys Asp Trp Ile  
 115 120 125  
 cca aat gcg gag ggc acg tcc ctt tac atc cgt ccg ttc atc atc gca 432  
 Pro Asn Ala Glu Gly Thr Ser Leu Tyr Ile Arg Pro Phe Ile Ile Ala  
 130 135 140  
 acc gag cct ttc ctt ggt gtt gcg gca tct cat acg tat aag ctc ttg 480  
 Thr Glu Pro Phe Leu Gly Val Ala Ala Ser His Thr Tyr Lys Leu Leu  
 145 150 155 160  
 atc att ctt tct ccg gtc ggc tct tat tac aaa gaa ggc att aag ccg 528  
 Ile Ile Leu Ser Pro Val Gly Ser Tyr Tyr Lys Glu Gly Ile Lys Pro

- 62 -

				165					170					175			
gtc	aaa	atc	gct	gtt	gaa	agt	gaa	ttt	gtc	cgt	gcg	gta	aaa	ggc	gga		576
Val	Lys	Ile	Ala	Val	Glu	Ser	Glu	Phe	Val	Arg	Ala	Val	Lys	Gly	Gly		
			180					185					190				
aca	gga	aat	gcc	aaa	acc	gca	gga	aac	tat	gct	tca	agc	tta	aaa	gcg		624
Thr	Gly	Asn	Ala	Lys	Thr	Ala	Gly	Asn	Tyr	Ala	Ser	Ser	Leu	Lys	Ala		
			195				200					205					
cag	cag	gta	gcc	gaa	gag	aaa	gga	ttt	tct	caa	gta	ctc	tg	ctg	gac		672
Gln	Gln	Val	Ala	Glu	Glu	Lys	Gly	Phe	Ser	Gln	Val	Leu	Trp	Leu	Asp		
	210					215					220						
ggc	att	gag	aag	aaa	tac	atc	gaa	gaa	gtc	gga	agc	atg	aac	atc	ttc		720
Gly	Ile	Glu	Lys	Lys	Tyr	Ile	Glu	Glu	Val	Gly	Ser	Met	Asn	Ile	Phe		
	225				230					235					240		
ttc	aaa	atc	aac	ggt	gaa	atc	gta	aca	ccg	atg	ctg	aac	ggg	agc	atc		768
Phe	Lys	Ile	Asn	Gly	Glu	Ile	Val	Thr	Pro	Met	Leu	Asn	Gly	Ser	Ile		
			245						250				255				
ctg	gaa	ggc	att	acg	cgc	aat	tca	gtc	atc	gcc	ttg	ctt	aag	cat	tg		816
Leu	Glu	Gly	Ile	Thr	Arg	Asn	Ser	Val	Ile	Ala	Leu	Leu	Lys	His	Trp		
			260					265					270				
ggc	ctt	caa	gtt	tca	gaa	cga	aaa	att	gcg	atc	gat	gag	gtc	atc	caa		864
Gly	Leu	Gln	Val	Ser	Glu	Arg	Lys	Ile	Ala	Ile	Asp	Glu	Val	Ile	Gln		
		275					280					285					
gcc	cat	aaa	gac	ggc	atc	ctg	gaa	gaa	gcc	ttc	gga	aca	ggt	aca	gca		912
Ala	His	Lys	Asp	Gly	Ile	Leu	Glu	Glu	Ala	Phe	Gly	Thr	Gly	Thr	Ala		
		290				295					300						
gct	gtt	att	tcc	cca	gtc	ggc	gag	ctg	atc	tg	cag	gat	gaa	aca	ctt		960
Ala	Val	Ile	Ser	Pro	Val	Gly	Glu	Leu	Ile	Trp	Gln	Asp	Glu	Thr	Leu		
	305				310					315					320		
tcg	atc	aac	aac	ggt	gaa	aca	gga	gaa	atc	gca	aaa	aaa	cta	tat	gac		1008
Ser	Ile	Asn	Asn	Gly	Glu	Thr	Gly	Glu	Ile	Ala	Lys	Lys	Leu	Tyr	Asp		
			325					330					335				
acg	att	aca	ggc	att	caa	aaa	ggc	gct	gtc	gca	gac	gaa	ttc	gga	tg		1056
Thr	Ile	Thr	Gly	Ile	Gln	Lys	Gly	Ala	Val	Ala	Asp	Glu	Phe	Gly	Trp		
			340				345					350					
acg	acc	gaa	gtc	gca	gcg	ctg	act	gaa	agc	aag	taa						1092
Thr	Thr	Glu	Val	Ala	Ala	Leu	Thr	Glu	Ser	Lys							
		355				360											

&lt;210&gt; 63

&lt;211&gt; 363

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 63

Met Thr Lys Gln Thr Ile Arg Val Glu Leu Thr Ser Thr Lys Lys Pro

- 63 -

1	5	10	15
Lys Pro Asp	Pro Asn Gln Leu Ser	Phe Gly Arg Val	Phe Thr Asp His
	20	25	30
Met Phe Val	Met Asp Tyr Ala	Ala Asp Lys Gly Trp	Tyr Asp Pro Arg
	35	40	45
Ile Ile Pro	Tyr Gln Pro Leu	Ser Met Asp Pro	Thr Ala Met Val Tyr
	50	55	60
His Tyr Gly	Gln Thr Val Phe	Glu Gly Leu Lys	Ala Tyr Val Ser Glu
	65	70	75
Asp Asp His	Val Leu Leu Phe	Arg Pro Glu Lys	Asn Met Glu Arg Leu
	85	90	95
Asn Gln Ser	Asn Asp Arg Leu	Cys Ile Pro Gln	Ile Asp Glu Glu Gln
	100	105	110
Val Leu Glu	Gly Leu Lys Gln	Leu Val Ala Ile	Asp Lys Asp Trp Ile
	115	120	125
Pro Asn Ala	Glu Gly Thr Ser	Leu Tyr Ile Arg	Pro Phe Ile Ile Ala
	130	135	140
Thr Glu Pro	Phe Leu Gly Val	Ala Ala Ser His	Thr Tyr Lys Leu Leu
	145	150	155
Ile Ile Leu	Ser Pro Val Gly	Ser Tyr Tyr Lys	Glu Gly Ile Lys Pro
	165	170	175
Val Lys Ile	Ala Val Glu Ser	Glu Phe Val Arg	Ala Val Lys Gly Gly
	180	185	190
Thr Gly Asn	Ala Lys Thr Ala	Gly Asn Tyr Ala	Ser Ser Leu Lys Ala
	195	200	205
Gln Gln Val	Ala Glu Glu Lys	Gly Phe Ser Gln	Val Leu Trp Leu Asp
	210	215	220
Gly Ile Glu	Lys Lys Tyr Ile	Glu Glu Val Gly	Ser Met Asn Ile Phe
	225	230	235
Phe Lys Ile	Asn Gly Glu Ile	Val Thr Pro Met	Leu Asn Gly Ser Ile
	245	250	255
Leu Glu Gly	Ile Thr Arg Asn	Ser Val Ile Ala	Leu Leu Lys His Trp
	260	265	270
Gly Leu Gln	Val Ser Glu Arg	Lys Ile Ala Ile	Asp Glu Val Ile Gln
	275	280	285
Ala His Lys	Asp Gly Ile Leu	Glu Glu Ala Phe	Gly Thr Gly Thr Ala
	290	295	300
Ala Val Ile	Ser Pro Val Gly	Glu Leu Ile Trp	Gln Asp Glu Thr Leu
	305	310	315
			320

<400> 64																
ttg	aat	aag	ctt	att	gaa	cga	gaa	aaa	act	gta	tat	tat	aag	gaa	aag	48
Met	Asn	Lys	Leu	Ile	Glu	Arg	Glu	Lys	Thr	Val	Tyr	Tyr	Lys	Glu	Lys	
1				5					10					15		
ccc	gac	ccg	tct	tcc	ttg	ggg	ttt	gga	caa	tat	ttt	aca	gat	tat	atg	96
Pro	Asp	Pro	Ser	Ser	Leu	Gly	Phe	Gly	Gln	Tyr	Phe	Thr	Asp	Tyr	Met	
			20					25					30			
ttt	gtg	atg	gac	tac	gaa	gag	ggg	att	gga	tgg	cat	cat	ccg	aga	att	144
Phe	Val	Met	Asp	Tyr	Glu	Glu	Gly	Ile	Gly	Trp	His	His	Pro	Arg	Ile	
		35					40					45				
gcg	ccg	tac	gca	ccg	ctt	acg	ctt	gat	ccg	tct	tca	tct	gtt	ttt	cat	192
Ala	Pro	Tyr	Ala	Pro	Leu	Thr	Leu	Asp	Pro	Ser	Ser	Ser	Val	Phe	His	
	50					55					60					
tac	ggc	cag	gct	gtt	ttt	gaa	gga	tta	aaa	gca	tac	aga	aca	gac	gac	240
Tyr	Gly	Gln	Ala	Val	Phe	Glu	Gly	Leu	Lys	Ala	Tyr	Arg	Thr	Asp	Asp	
65					70					75					80	
ggc	agg	gtg	ctg	ctg	ttc	cgt	ccg	gat	caa	aat	atc	aaa	cgg	ctg	aac	288
Gly	Arg	Val	Leu	Leu	Phe	Arg	Pro	Asp	Gln	Asn	Ile	Lys	Arg	Leu	Asn	
				85					90					95		
aga	tcg	tgt	gag	cgc	atg	agc	atg	ccc	cct	tta	gac	gaa	gag	ctg	gtg	336
Arg	Ser	Cys	Glu	Arg	Met	Ser	Met	Pro	Pro	Leu	Asp	Glu	Glu	Leu	Val	
			100					105						110		
ctt	gag	gca	ttg	acg	caa	tta	gtt	gag	ctg	gag	aaa	gat	tgg	gtt	cca	384
Leu	Glu	Ala	Leu	Thr	Gln	Leu	Val	Glu	Leu	Glu	Lys	Asp	Trp	Val	Pro	
		115					120					125				
aag	gaa	aaa	gga	acg	tca	ctg	tat	att	cgt	cct	ttt	gtc	att	gcc	aca	432
Lys	Glu	Lys	Gly	Thr	Ser	Leu	Tyr	Ile	Arg	Pro	Phe	Val	Ile	Ala	Thr	
	130					135					140					
gaa	ccg	agt	ctc	ggt	gtg	aag	gca	tcc	agg	agc	tat	aca	ttt	atg	atc	480

- 65 -

Glu	Pro	Ser	Leu	Gly	Val	Lys	Ala	Ser	Arg	Ser	Tyr	Thr	Phe	Met	Ile	
145					150					155					160	
gtg	ctt	tcg	cct	gtc	ggc	tcc	tat	tat	ggc	gac	gat	cag	ctg	aag	ccg	528
Val	Leu	Ser	Pro	Val	Gly	Ser	Tyr	Tyr	Gly	Asp	Asp	Gln	Leu	Lys	Pro	
				165					170					175		
gtt	aga	atc	tat	gtc	gaa	gat	gag	tat	gtg	agg	gcg	gtc	aac	gga	gga	576
Val	Arg	Ile	Tyr	Val	Glu	Asp	Glu	Tyr	Val	Arg	Ala	Val	Asn	Gly	Gly	
			180					185					190			
gtc	ggg	ttt	gca	aaa	acg	gct	gga	aac	tat	gcc	gcc	agt	ctt	cag	gca	624
Val	Gly	Phe	Ala	Lys	Thr	Ala	Gly	Asn	Tyr	Ala	Ala	Ser	Leu	Gln	Ala	
		195					200					205				
cag	cgg	aaa	gcg	aat	gaa	ctg	ggc	tat	gac	cag	gta	ctg	tgg	ctg	gac	672
Gln	Arg	Lys	Ala	Asn	Glu	Leu	Gly	Tyr	Asp	Gln	Val	Leu	Trp	Leu	Asp	
	210					215					220					
gcc	atc	gaa	aag	aaa	tat	gtg	gaa	gaa	gta	ggg	agc	atg	aac	atc	ttt	720
Ala	Ile	Glu	Lys	Lys	Tyr	Val	Glu	Glu	Val	Gly	Ser	Met	Asn	Ile	Phe	
225					230					235					240	
ttc	gtc	ata	aac	ggg	gaa	gct	gtc	aca	cct	gct	tta	agc	gga	agc	att	768
Phe	Val	Ile	Asn	Gly	Glu	Ala	Val	Thr	Pro	Ala	Leu	Ser	Gly	Ser	Ile	
			245						250					255		
tta	agc	ggg	gtt	aca	cgt	gcg	tct	gcg	att	gaa	ttg	att	cga	agc	tgg	816
Leu	Ser	Gly	Val	Thr	Arg	Ala	Ser	Ala	Ile	Glu	Leu	Ile	Arg	Ser	Trp	
			260					265					270			
ggc	att	ccg	gtt	cgt	gaa	gag	aga	ata	tcg	att	gat	gag	gtg	tat	gcg	864
Gly	Ile	Pro	Val	Arg	Glu	Glu	Arg	Ile	Ser	Ile	Asp	Glu	Val	Tyr	Ala	
		275					280					285				
gcc	tct	gca	cgc	gga	gaa	ttg	aca	gag	gtc	ttt	ggc	aca	ggc	acg	gca	912
Ala	Ser	Ala	Arg	Gly	Glu	Leu	Thr	Glu	Val	Phe	Gly	Thr	Gly	Thr	Ala	
	290					295					300					
gca	gtc	gtt	acg	cct	gtc	ggt	gaa	ctc	aac	atc	cat	gga	aaa	acg	gtg	960
Ala	Val	Val	Thr	Pro	Val	Gly	Glu	Leu	Asn	Ile	His	Gly	Lys	Thr	Val	
305					310					315					320	
att	gta	ggc	gac	ggg	caa	atc	ggg	gac	ctc	tcg	aaa	aag	ctg	tat	gaa	1008
Ile	Val	Gly	Asp	Gly	Gln	Ile	Gly	Asp	Leu	Ser	Lys	Lys	Leu	Tyr	Glu	
				325				330						335		
acg	ata	aca	gat	att	cag	ctt	ggc	aag	gta	aaa	ggc	ccg	ttt	aac	tgg	1056
Thr	Ile	Thr	Asp	Ile	Gln	Leu	Gly	Lys	Val	Lys	Gly	Pro	Phe	Asn	Trp	
			340					345					350			
aca	gtg	gaa	gtg	tga												1071
Thr	Val	Glu	Val													
			355													

&lt;210&gt; 65

&lt;211&gt; 356



- 66 -

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 65

Met	Asn	Lys	Leu	Ile	Glu	Arg	Glu	Lys	Thr	Val	Tyr	Tyr	Lys	Glu	Lys	1	5	10	15
Pro	Asp	Pro	Ser	Ser	Leu	Gly	Phe	Gly	Gln	Tyr	Phe	Thr	Asp	Tyr	Met	20	25	30	
Phe	Val	Met	Asp	Tyr	Glu	Glu	Gly	Ile	Gly	Trp	His	His	Pro	Arg	Ile	35	40	45	
Ala	Pro	Tyr	Ala	Pro	Leu	Thr	Leu	Asp	Pro	Ser	Ser	Ser	Val	Phe	His	50	55	60	
Tyr	Gly	Gln	Ala	Val	Phe	Glu	Gly	Leu	Lys	Ala	Tyr	Arg	Thr	Asp	Asp	65	70	75	80
Gly	Arg	Val	Leu	Leu	Phe	Arg	Pro	Asp	Gln	Asn	Ile	Lys	Arg	Leu	Asn	85	90	95	
Arg	Ser	Cys	Glu	Arg	Met	Ser	Met	Pro	Pro	Leu	Asp	Glu	Glu	Leu	Val	100	105	110	
Leu	Glu	Ala	Leu	Thr	Gln	Leu	Val	Glu	Leu	Glu	Lys	Asp	Trp	Val	Pro	115	120	125	
Lys	Glu	Lys	Gly	Thr	Ser	Leu	Tyr	Ile	Arg	Pro	Phe	Val	Ile	Ala	Thr	130	135	140	
Glu	Pro	Ser	Leu	Gly	Val	Lys	Ala	Ser	Arg	Ser	Tyr	Thr	Phe	Met	Ile	145	150	155	160
Val	Leu	Ser	Pro	Val	Gly	Ser	Tyr	Tyr	Gly	Asp	Asp	Gln	Leu	Lys	Pro	165	170	175	
Val	Arg	Ile	Tyr	Val	Glu	Asp	Glu	Tyr	Val	Arg	Ala	Val	Asn	Gly	Gly	180	185	190	
Val	Gly	Phe	Ala	Lys	Thr	Ala	Gly	Asn	Tyr	Ala	Ala	Ser	Leu	Gln	Ala	195	200	205	
Gln	Arg	Lys	Ala	Asn	Glu	Leu	Gly	Tyr	Asp	Gln	Val	Leu	Trp	Leu	Asp	210	215	220	
Ala	Ile	Glu	Lys	Lys	Tyr	Val	Glu	Glu	Val	Gly	Ser	Met	Asn	Ile	Phe	225	230	235	240
Phe	Val	Ile	Asn	Gly	Glu	Ala	Val	Thr	Pro	Ala	Leu	Ser	Gly	Ser	Ile	245	250	255	
Leu	Ser	Gly	Val	Thr	Arg	Ala	Ser	Ala	Ile	Glu	Leu	Ile	Arg	Ser	Trp	260	265	270	
Gly	Ile	Pro	Val	Arg	Glu	Glu	Arg	Ile	Ser	Ile	Asp	Glu	Val	Tyr	Ala	275	280	285	

- 67 -

Ala Ser Ala Arg Gly Glu Leu Thr Glu Val Phe Gly Thr Gly Thr Ala  
 290 295 300

Ala Val Val Thr Pro Val Gly Glu Leu Asn Ile His Gly Lys Thr Val  
 305 310 315 320

Ile Val Gly Asp Gly Gln Ile Gly Asp Leu Ser Lys Lys Leu Tyr Glu  
 325 330 335

Thr Ile Thr Asp Ile Gln Leu Gly Lys Val Lys Gly Pro Phe Asn Trp  
 340 345 350

Thr Val Glu Val  
 355

&lt;210&gt; 66

&lt;211&gt; 1428

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1425)

&lt;400&gt; 66

atg tta aac ggc caa aaa gaa tat cgc gtg gaa aaa gac ttc ctt ggg	48
Met Leu Asn Gly Gln Lys Glu Tyr Arg Val Glu Lys Asp Phe Leu Gly	
1 5 10 15	
gaa aaa caa att gaa gca gat gtt tat tac gga att cag acg ctc cgt	96
Glu Lys Gln Ile Glu Ala Asp Val Tyr Tyr Gly Ile Gln Thr Leu Arg	
20 25 30	
gct tct gaa aat ttt ccg atc aca gga tac aaa atc cat gag gaa atg	144
Ala Ser Glu Asn Phe Pro Ile Thr Gly Tyr Lys Ile His Glu Glu Met	
35 40 45	
att aac gca ctg gcg att gtg aaa aaa gct gcg gct ctt gcc aac atg	192
Ile Asn Ala Leu Ala Ile Val Lys Lys Ala Ala Ala Leu Ala Asn Met	
50 55 60	
gac gtg aaa cgg ctg tat gaa gga att ggc caa gct atc gta caa gcc	240
Asp Val Lys Arg Leu Tyr Glu Gly Ile Gly Gln Ala Ile Val Gln Ala	
65 70 75 80	
gct gac gag att ctg gaa ggc aag tgg cac gat cag ttt atc gtc gat	288
Ala Asp Glu Ile Leu Glu Gly Lys Trp His Asp Gln Phe Ile Val Asp	
85 90 95	
ccg att cag ggc ggt gcc gga act tct atg aac atg aac gcg aat gag	336
Pro Ile Gln Gly Gly Ala Gly Thr Ser Met Asn Met Asn Ala Asn Glu	
100 105 110	
gtt atc gga aac cgg gcg ctt gaa atc atg gga cat aaa aag gga gat	384
Val Ile Gly Asn Arg Ala Leu Glu Ile Met Gly His Lys Lys Gly Asp	
115 120 125	

- 68 -

tat atc cat tta agt cca aac aca cat gtg aac atg tca cag tct cag	432
Tyr Ile His Leu Ser Pro Asn Thr His Val Asn Met Ser Gln Ser Gln	
130 135 140	
aac gat gtg ttc ccg act gct atc cat att tcc aca ttg aag ctc tta	480
Asn Asp Val Phe Pro Thr Ala Ile His Ile Ser Thr Leu Lys Leu Leu	
145 150 155 160	
gaa aaa ctg ctg aaa aca atg gaa gat atg cat agt gtg ttt aaa caa	528
Glu Lys Leu Leu Lys Thr Met Glu Asp Met His Ser Val Phe Lys Gln	
165 170 175	
aaa gca cag gag ttt cac tct gtt att aaa atg ggc cgg aca cac ctt	576
Lys Ala Gln Glu Phe His Ser Val Ile Lys Met Gly Arg Thr His Leu	
180 185 190	
caa gat gcg gtt ccg atc cgt ctt ggc cag gaa ttc gaa gct tac agc	624
Gln Asp Ala Val Pro Ile Arg Leu Gly Gln Glu Phe Glu Ala Tyr Ser	
195 200 205	
cgt gtt ctc gag cgt gat atc aaa cga atc aag caa tcg cgc cag cac	672
Arg Val Leu Glu Arg Asp Ile Lys Arg Ile Lys Gln Ser Arg Gln His	
210 215 220	
ctg tat gaa gtc aac atg ggc gca act gct gtt ggt aca ggg ctg aac	720
Leu Tyr Glu Val Asn Met Gly Ala Thr Ala Val Gly Thr Gly Leu Asn	
225 230 235 240	
gct gat cct gaa tat atc aaa cag gta gta aag cac ctt gct gat att	768
Ala Asp Pro Glu Tyr Ile Lys Gln Val Val Lys His Leu Ala Asp Ile	
245 250 255	
agc ggg ctt cct ctt gtc ggc gct gat cat ctt gtt gat gcg aca caa	816
Ser Gly Leu Pro Leu Val Gly Ala Asp His Leu Val Asp Ala Thr Gln	
260 265 270	
aat aca gat gcc tat aca gag gta tca gct tca tta aaa gtc tgc atg	864
Asn Thr Asp Ala Tyr Thr Glu Val Ser Ala Ser Leu Lys Val Cys Met	
275 280 285	
atg aac atg tcg aag atc gca aac gac ctg cgc tta atg gcg tcg gga	912
Met Asn Met Ser Lys Ile Ala Asn Asp Leu Arg Leu Met Ala Ser Gly	
290 295 300	
ccg cgc gcc gga ctt gcg gaa att tct ctg cct gca cgt cag ccg ggt	960
Pro Arg Ala Gly Leu Ala Glu Ile Ser Leu Pro Ala Arg Gln Pro Gly	
305 310 315 320	
tca tct att atg ccg ggg aaa gtc aat ccg gtt atg gcg gag ctg atc	1008
Ser Ser Ile Met Pro Gly Lys Val Asn Pro Val Met Ala Glu Leu Ile	
325 330 335	
aac caa att gcg ttc cag gtt atc gga aat gac aat aca atc tgc ctt	1056
Asn Gln Ile Ala Phe Gln Val Ile Gly Asn Asp Asn Thr Ile Cys Leu	
340 345 350	
gct tca gaa gcc ggc cag ctt gag ttg aac gtc atg gag ccc gtg ctt	1104
Ala Ser Glu Ala Gly Gln Leu Glu Leu Asn Val Met Glu Pro Val Leu	

- 69 -

355					360					365						
gtc	ttt	aat	ttg	ctt	caa	tcc	atc	agc	atc	atg	aac	aac	ggc	ttc	cgt	1152
Val	Phe	Asn	Leu	Leu	Gln	Ser	Ile	Ser	Ile	Met	Asn	Asn	Gly	Phe	Arg	
370					375					380						
tcg	ttc	act	gac	aac	tgc	tta	aaa	ggc	att	gaa	gcc	aac	gaa	aag	cgt	1200
Ser	Phe	Thr	Asp	Asn	Cys	Leu	Lys	Gly	Ile	Glu	Ala	Asn	Glu	Lys	Arg	
385					390					395					400	
atg	aag	caa	tac	gta	gaa	aaa	agc	gca	ggc	gtg	atc	aca	gct	gtc	aat	1248
Met	Lys	Gln	Tyr	Val	Glu	Lys	Ser	Ala	Gly	Val	Ile	Thr	Ala	Val	Asn	
405					410					415						
ccg	cat	ctt	ggg	tat	gaa	gcg	gca	gct	aga	att	gcc	agg	gaa	gca	att	1296
Pro	His	Leu	Gly	Tyr	Glu	Ala	Ala	Ala	Arg	Ile	Ala	Arg	Glu	Ala	Ile	
420					425					430						
atg	aca	ggg	caa	tct	gtc	cgg	gat	ctt	tgt	ctg	cag	cat	gat	gtg	ctg	1344
Met	Thr	Gly	Gln	Ser	Val	Arg	Asp	Leu	Cys	Leu	Gln	His	Asp	Val	Leu	
435					440					445						
act	gaa	gaa	gaa	ttg	gat	att	att	tta	aac	cca	tat	gag	atg	acc	aaa	1392
Thr	Glu	Glu	Glu	Leu	Asp	Ile	Ile	Leu	Asn	Pro	Tyr	Glu	Met	Thr	Lys	
450					455					460						
cca	ggt	atc	gca	ggg	aaa	gaa	cta	tta	gaa	aaa	taa					1428
Pro	Gly	Ile	Ala	Gly	Lys	Glu	Leu	Leu	Glu	Lys						
465					470					475						

&lt;210&gt; 67

&lt;211&gt; 475

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 67

Met	Leu	Asn	Gly	Gln	Lys	Glu	Tyr	Arg	Val	Glu	Lys	Asp	Phe	Leu	Gly	
1				5					10					15		
Glu	Lys	Gln	Ile	Glu	Ala	Asp	Val	Tyr	Tyr	Gly	Ile	Gln	Thr	Leu	Arg	
		20						25					30			
Ala	Ser	Glu	Asn	Phe	Pro	Ile	Thr	Gly	Tyr	Lys	Ile	His	Glu	Glu	Met	
		35					40					45				
Ile	Asn	Ala	Leu	Ala	Ile	Val	Lys	Lys	Ala	Ala	Ala	Leu	Ala	Asn	Met	
	50				55						60					
Asp	Val	Lys	Arg	Leu	Tyr	Glu	Gly	Ile	Gly	Gln	Ala	Ile	Val	Gln	Ala	
65				70						75				80		
Ala	Asp	Glu	Ile	Leu	Glu	Gly	Lys	Trp	His	Asp	Gln	Phe	Ile	Val	Asp	
			85					90					95			
Pro	Ile	Gln	Gly	Gly	Ala	Gly	Thr	Ser	Met	Asn	Met	Asn	Ala	Asn	Glu	
		100						105					110			

- 70 -

Val Ile Gly Asn Arg Ala Leu Glu Ile Met Gly His Lys Lys Gly Asp  
 115 120 125  
 Tyr Ile His Leu Ser Pro Asn Thr His Val Asn Met Ser Gln Ser Gln  
 130 135 140  
 Asn Asp Val Phe Pro Thr Ala Ile His Ile Ser Thr Leu Lys Leu Leu  
 145 150 155 160  
 Glu Lys Leu Leu Lys Thr Met Glu Asp Met His Ser Val Phe Lys Gln  
 165 170 175  
 Lys Ala Gln Glu Phe His Ser Val Ile Lys Met Gly Arg Thr His Leu  
 180 185 190  
 Gln Asp Ala Val Pro Ile Arg Leu Gly Gln Glu Phe Glu Ala Tyr Ser  
 195 200 205  
 Arg Val Leu Glu Arg Asp Ile Lys Arg Ile Lys Gln Ser Arg Gln His  
 210 215 220  
 Leu Tyr Glu Val Asn Met Gly Ala Thr Ala Val Gly Thr Gly Leu Asn  
 225 230 235 240  
 Ala Asp Pro Glu Tyr Ile Lys Gln Val Val Lys His Leu Ala Asp Ile  
 245 250 255  
 Ser Gly Leu Pro Leu Val Gly Ala Asp His Leu Val Asp Ala Thr Gln  
 260 265 270  
 Asn Thr Asp Ala Tyr Thr Glu Val Ser Ala Ser Leu Lys Val Cys Met  
 275 280 285  
 Met Asn Met Ser Lys Ile Ala Asn Asp Leu Arg Leu Met Ala Ser Gly  
 290 295 300  
 Pro Arg Ala Gly Leu Ala Glu Ile Ser Leu Pro Ala Arg Gln Pro Gly  
 305 310 315 320  
 Ser Ser Ile Met Pro Gly Lys Val Asn Pro Val Met Ala Glu Leu Ile  
 325 330 335  
 Asn Gln Ile Ala Phe Gln Val Ile Gly Asn Asp Asn Thr Ile Cys Leu  
 340 345 350  
 Ala Ser Glu Ala Gly Gln Leu Glu Leu Asn Val Met Glu Pro Val Leu  
 355 360 365  
 Val Phe Asn Leu Leu Gln Ser Ile Ser Ile Met Asn Asn Gly Phe Arg  
 370 375 380  
 Ser Phe Thr Asp Asn Cys Leu Lys Gly Ile Glu Ala Asn Glu Lys Arg  
 385 390 395 400  
 Met Lys Gln Tyr Val Glu Lys Ser Ala Gly Val Ile Thr Ala Val Asn  
 405 410 415  
 Pro His Leu Gly Tyr Glu Ala Ala Ala Arg Ile Ala Arg Glu Ala Ile

- 71 -

420                      425                      430  
 Met Thr Gly Gln Ser Val Arg Asp Leu Cys Leu Gln His Asp Val Leu  
 435                      440                      445  
 Thr Glu Glu Glu Leu Asp Ile Ile Leu Asn Pro Tyr Glu Met Thr Lys  
 450                      455                      460  
 Pro Gly Ile Ala Gly Lys Glu Leu Leu Glu Lys  
 465                      470                      475

&lt;210&gt; 68

&lt;211&gt; 768

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(765)

&lt;400&gt; 68

atg aaa cga gaa agc aac att caa gtg ctc agc cgt ggt caa aaa gat 48  
 Met Lys Arg Glu Ser Asn Ile Gln Val Leu Ser Arg Gly Gln Lys Asp  
 1                      5                      10                      15

cag cct gtg agc cag att tat caa gta tca aca atg act tct cta tta 96  
 Gln Pro Val Ser Gln Ile Tyr Gln Val Ser Thr Met Thr Ser Leu Leu  
 20                      25                      30

gac gga gta tat gac gga gat ttt gaa ctg tca gag att ccg aaa tat 144  
 Asp Gly Val Tyr Asp Gly Asp Phe Glu Leu Ser Glu Ile Pro Lys Tyr  
 35                      40                      45

gga gac ttc ggt atc gga acc ttt aac aag ctt gac gga gag ctg att 192  
 Gly Asp Phe Gly Ile Gly Thr Phe Asn Lys Leu Asp Gly Glu Leu Ile  
 50                      55                      60

ggg ttt gac ggc gaa ttt tac cgt ctt cgc tca gac gga acc gcg aca 240  
 Gly Phe Asp Gly Glu Phe Tyr Arg Leu Arg Ser Asp Gly Thr Ala Thr  
 65                      70                      75                      80

ccg gtc caa aat gga gac cgt tca ccg ttc tgt tca ttt acg ttc ttt 288  
 Pro Val Gln Asn Gly Asp Arg Ser Pro Phe Cys Ser Phe Thr Phe Phe  
 85                      90                      95

aca ccg gac atg acg cac aaa att gat gcg aaa atg aca cgc gaa gac 336  
 Thr Pro Asp Met Thr His Lys Ile Asp Ala Lys Met Thr Arg Glu Asp  
 100                      105                      110

ttt gaa aaa gag atc aac agc atg ctg cca agc aga aac tta ttt tat 384  
 Phe Glu Lys Glu Ile Asn Ser Met Leu Pro Ser Arg Asn Leu Phe Tyr  
 115                      120                      125

gca att cgc att gac gga ttg ttt aaa aag gtg cag aca aga aca gta 432  
 Ala Ile Arg Ile Asp Gly Leu Phe Lys Lys Val Gln Thr Arg Thr Val  
 130                      135                      140

- 72 -

gaa ctt caa gaa aaa cct tac gtg cca atg gtt gaa gcg gtc aaa aca 480  
 Glu Leu Gln Glu Lys Pro Tyr Val Pro Met Val Glu Ala Val Lys Thr  
 145 150 155 160

cag ccg att ttc aac ttc gac aac gtg aga gga acg att gta ggt ttc 528  
 Gln Pro Ile Phe Asn Phe Asp Asn Val Arg Gly Thr Ile Val Gly Phe  
 165 170 175

ttg aca cca gct tat gca aac gga atc gcc gtt tct ggc tat cac ctg 576  
 Leu Thr Pro Ala Tyr Ala Asn Gly Ile Ala Val Ser Gly Tyr His Leu  
 180 185 190

cac ttc att gac gaa gga cgc aat tca ggc gga cac gtt ttt gac tat 624  
 His Phe Ile Asp Glu Gly Arg Asn Ser Gly Gly His Val Phe Asp Tyr  
 195 200 205

gtg ctt gag gat tgc acg gtt acg att tct caa aaa atg aac atg aat 672  
 Val Leu Glu Asp Cys Thr Val Thr Ile Ser Gln Lys Met Asn Met Asn  
 210 215 220

ctc aga ctt ccg aac aca gcg gat ttc ttt aat gcg aat ctg gat aac 720  
 Leu Arg Leu Pro Asn Thr Ala Asp Phe Phe Asn Ala Asn Leu Asp Asn  
 225 230 235 240

cct gat ttt gcg aaa gat atc gaa aca act gaa gga agc cct gaa taa 768  
 Pro Asp Phe Ala Lys Asp Ile Glu Thr Thr Glu Gly Ser Pro Glu  
 245 250 255

&lt;210&gt; 69

&lt;211&gt; 255

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 69

Met Lys Arg Glu Ser Asn Ile Gln Val Leu Ser Arg Gly Gln Lys Asp  
 1 5 10 15

Gln Pro Val Ser Gln Ile Tyr Gln Val Ser Thr Met Thr Ser Leu Leu  
 20 25 30

Asp Gly Val Tyr Asp Gly Asp Phe Glu Leu Ser Glu Ile Pro Lys Tyr  
 35 40 45

Gly Asp Phe Gly Ile Gly Thr Phe Asn Lys Leu Asp Gly Glu Leu Ile  
 50 55 60

Gly Phe Asp Gly Glu Phe Tyr Arg Leu Arg Ser Asp Gly Thr Ala Thr  
 65 70 75 80

Pro Val Gln Asn Gly Asp Arg Ser Pro Phe Cys Ser Phe Thr Phe Phe  
 85 90 95

Thr Pro Asp Met Thr His Lys Ile Asp Ala Lys Met Thr Arg Glu Asp  
 100 105 110

Phe Glu Lys Glu Ile Asn Ser Met Leu Pro Ser Arg Asn Leu Phe Tyr  
 115 120 125

- 73 -

Ala Ile Arg Ile Asp Gly Leu Phe Lys Lys Val Gln Thr Arg Thr Val  
130 135 140

Glu Leu Gln Glu Lys Pro Tyr Val Pro Met Val Glu Ala Val Lys Thr  
145 150 155 160

Gln Pro Ile Phe Asn Phe Asp Asn Val Arg Gly Thr Ile Val Gly Phe  
165 170 175

Leu Thr Pro Ala Tyr Ala Asn Gly Ile Ala Val Ser Gly Tyr His Leu  
180 185 190

His Phe Ile Asp Glu Gly Arg Asn Ser Gly Gly His Val Phe Asp Tyr  
195 200 205

Val Leu Glu Asp Cys Thr Val Thr Ile Ser Gln Lys Met Asn Met Asn  
210 215 220

Leu Arg Leu Pro Asn Thr Ala Asp Phe Phe Asn Ala Asn Leu Asp Asn  
225 230 235 240

Pro Asp Phe Ala Lys Asp Ile Glu Thr Thr Glu Gly Ser Pro Glu  
245 250 255

<210> 70

<211> 1254

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (1)..(1251)

<400> 70

atg aca ttc tcc ctt ttt ggt gac aaa ttt acc cgc cac tcc ggc att 48  
Met Thr Phe Ser Leu Phe Gly Asp Lys Phe Thr Arg His Ser Gly Ile  
1 5 10 15

acg ctg ttg atg gaa gat ctg aac gac ggt tta cgc acg cct ggc gcg 96  
Thr Leu Leu Met Glu Asp Leu Asn Asp Gly Leu Arg Thr Pro Gly Ala  
20 25 30

att atg ctc ggc ggc ggt aat ccg gcg cag atc ccg gaa atg cag gac 144  
Ile Met Leu Gly Gly Gly Asn Pro Ala Gln Ile Pro Glu Met Gln Asp  
35 40 45

tac ttc cag acg cta ctg acc gac atg ctg gaa agt ggc aaa gcg act 192  
Tyr Phe Gln Thr Leu Leu Thr Asp Met Leu Glu Ser Gly Lys Ala Thr  
50 55 60

gat gca ctg tgt aac tac gac ggt cca cag ggg aaa acg gag cta ctc 240  
Asp Ala Leu Cys Asn Tyr Asp Gly Pro Gln Gly Lys Thr Glu Leu Leu  
65 70 75 80

aca ctg ctt gcc gga atg ctg cgc gag aag ttg ggt tgg gat atc gaa 288  
Thr Leu Leu Ala Gly Met Leu Arg Glu Lys Leu Gly Trp Asp Ile Glu



- 74 -

85								90				95				
cca	cag	aat	att	gca	cta	aca	aac	ggc	agc	cag	agc	gcg	ttt	ttc	tac	336
Pro	Gln	Asn	Ile	Ala	Leu	Thr	Asn	Gly	Ser	Gln	Ser	Ala	Phe	Phe	Tyr	
			100					105					110			
tta	ttt	aac	ctg	ttt	gcc	gga	cgc	cgt	gcc	gat	ggt	cgg	gtc	aaa	aaa	384
Leu	Phe	Asn	Leu	Phe	Ala	Gly	Arg	Arg	Ala	Asp	Gly	Arg	Val	Lys	Lys	
		115					120					125				
gtg	ctg	ttc	ccg	ctt	gca	ccg	gaa	tac	att	ggc	tat	gct	gac	gcc	gga	432
Val	Leu	Phe	Pro	Leu	Ala	Pro	Glu	Tyr	Ile	Gly	Tyr	Ala	Asp	Ala	Gly	
	130					135					140					
ctg	gaa	gaa	gat	ctg	ttt	gtc	tct	gcg	cgt	ccg	aat	att	gaa	ctg	ctg	480
Leu	Glu	Glu	Asp	Leu	Phe	Val	Ser	Ala	Arg	Pro	Asn	Ile	Glu	Leu	Leu	
145					150					155					160	
ccg	gaa	ggc	cag	ttt	aaa	tac	cac	gtc	gat	ttt	gag	cat	ctg	cat	att	528
Pro	Glu	Gly	Gln	Phe	Lys	Tyr	His	Val	Asp	Phe	Glu	His	Leu	His	Ile	
			165					170					175			
ggc	gaa	gaa	acc	ggg	atg	att	tgc	gtc	tcc	cgg	ccg	acg	aat	cca	aca	576
Gly	Glu	Glu	Thr	Gly	Met	Ile	Cys	Val	Ser	Arg	Pro	Thr	Asn	Pro	Thr	
			180				185						190			
ggc	aat	gtg	att	act	gac	gaa	gag	ttg	ctg	aag	ctt	gac	gcg	ctg	ggc	624
Gly	Asn	Val	Ile	Thr	Asp	Glu	Glu	Leu	Leu	Lys	Leu	Asp	Ala	Leu	Gly	
		195					200					205				
aat	caa	cac	ggc	att	ccg	ctg	gtg	att	gat	aac	gct	tat	ggc	gtc	ccg	672
Asn	Gln	His	Gly	Ile	Pro	Leu	Val	Ile	Asp	Asn	Ala	Tyr	Gly	Val	Pro	
	210					215					220					
ttc	ccg	ggg	atc	atc	ttc	agt	gaa	gcg	cgc	ccg	cta	tgg	aat	ccg	aat	720
Phe	Pro	Gly	Ile	Ile	Phe	Ser	Glu	Ala	Arg	Pro	Leu	Trp	Asn	Pro	Asn	
225					230				235					240		
atc	gtg	ctg	tgc	atg	agt	ctt	tcc	aag	ctg	ggg	cta	cct	ggc	tcc	cgc	768
Ile	Val	Leu	Cys	Met	Ser	Leu	Ser	Lys	Leu	Gly	Leu	Pro	Gly	Ser	Arg	
			245					250					255			
tgc	ggc	att	atc	atc	gcc	aat	gaa	aaa	atc	atc	acc	gcc	atc	acc	aat	816
Cys	Gly	Ile	Ile	Ile	Ala	Asn	Glu	Lys	Ile	Ile	Thr	Ala	Ile	Thr	Asn	
			260				265					270				
atg	aac	ggc	att	atc	agc	ctg	gca	cct	ggc	ggg	att	ggg	ccg	gcg	atg	864
Met	Asn	Gly	Ile	Ile	Ser	Leu	Ala	Pro	Gly	Gly	Ile	Gly	Pro	Ala	Met	
		275					280					285				
atg	tgt	gaa	atg	att	aag	cgt	aac	gat	ctg	ctg	cgc	ctg	tct	gaa	aca	912
Met	Cys	Glu	Met	Ile	Lys	Arg	Asn	Asp	Leu	Leu	Arg	Leu	Ser	Glu	Thr	
	290					295			300							
gtc	atc	aaa	ccg	ttt	tac	tac	cag	cgt	gtt	cag	gaa	act	atc	gcc	atc	960
Val	Ile	Lys	Pro	Phe	Tyr	Tyr	Gln	Arg	Val	Gln	Glu	Thr	Ile	Ala	Ile	
305				310					315						320	

- 75 -

att cgc cgc tat tta ccg gaa aat cgc tgc ctg att cat aaa ccg gaa 1008  
 Ile Arg Arg Tyr Leu Pro Glu Asn Arg Cys Leu Ile His Lys Pro Glu  
 325 330 335

gga gcc att ttc ctc tgg cta tgg ttt aag gat ttg ccc att acg acc 1056  
 Gly Ala Ile Phe Leu Trp Leu Trp Phe Lys Asp Leu Pro Ile Thr Thr  
 340 345 350

aag cag ctc tat cag cgc ctg aaa gca cgc ggc gtg ctg atg gtg ccg 1104  
 Lys Gln Leu Tyr Gln Arg Leu Lys Ala Arg Gly Val Leu Met Val Pro  
 355 360 365

ggg cac aac ttc ttc cca ggg ctg gat aaa ccg tgg ccg cat acg cat 1152  
 Gly His Asn Phe Phe Pro Gly Leu Asp Lys Pro Trp Pro His Thr His  
 370 375 380

caa tgt atg cgc atg aac tac gta cca gag ccg gag aaa att gag gcg 1200  
 Gln Cys Met Arg Met Asn Tyr Val Pro Glu Pro Glu Lys Ile Glu Ala  
 385 390 395 400

ggg gtg aag att ctg gcg gaa gag ata gaa aga gcc tgg gct gaa agt 1248  
 Gly Val Lys Ile Leu Ala Glu Glu Ile Glu Arg Ala Trp Ala Glu Ser  
 405 410 415

cac taa 1254  
 His

&lt;210&gt; 71

&lt;211&gt; 417

&lt;212&gt; PRT

&lt;213&gt; Escherichia coli

&lt;400&gt; 71

Met Thr Phe Ser Leu Phe Gly Asp Lys Phe Thr Arg His Ser Gly Ile  
 1 5 10 15

Thr Leu Leu Met Glu Asp Leu Asn Asp Gly Leu Arg Thr Pro Gly Ala  
 20 25 30

Ile Met Leu Gly Gly Gly Asn Pro Ala Gln Ile Pro Glu Met Gln Asp  
 35 40 45

Tyr Phe Gln Thr Leu Leu Thr Asp Met Leu Glu Ser Gly Lys Ala Thr  
 50 55 60

Asp Ala Leu Cys Asn Tyr Asp Gly Pro Gln Gly Lys Thr Glu Leu Leu  
 65 70 75 80

Thr Leu Leu Ala Gly Met Leu Arg Glu Lys Leu Gly Trp Asp Ile Glu  
 85 90 95

Pro Gln Asn Ile Ala Leu Thr Asn Gly Ser Gln Ser Ala Phe Phe Tyr  
 100 105 110

Leu Phe Asn Leu Phe Ala Gly Arg Arg Ala Asp Gly Arg Val Lys Lys  
 115 120 125

- 76 -

Val Leu Phe Pro Leu Ala Pro Glu Tyr Ile Gly Tyr Ala Asp Ala Gly  
 130 135 140  
 Leu Glu Glu Asp Leu Phe Val Ser Ala Arg Pro Asn Ile Glu Leu Leu  
 145 150 155 160  
 Pro Glu Gly Gln Phe Lys Tyr His Val Asp Phe Glu His Leu His Ile  
 165 170 175  
 Gly Glu Glu Thr Gly Met Ile Cys Val Ser Arg Pro Thr Asn Pro Thr  
 180 185 190  
 Gly Asn Val Ile Thr Asp Glu Glu Leu Leu Lys Leu Asp Ala Leu Gly  
 195 200 205  
 Asn Gln His Gly Ile Pro Leu Val Ile Asp Asn Ala Tyr Gly Val Pro  
 210 215 220  
 Phe Pro Gly Ile Ile Phe Ser Glu Ala Arg Pro Leu Trp Asn Pro Asn  
 225 230 235 240  
 Ile Val Leu Cys Met Ser Leu Ser Lys Leu Gly Leu Pro Gly Ser Arg  
 245 250 255  
 Cys Gly Ile Ile Ile Ala Asn Glu Lys Ile Ile Thr Ala Ile Thr Asn  
 260 265 270  
 Met Asn Gly Ile Ile Ser Leu Ala Pro Gly Gly Ile Gly Pro Ala Met  
 275 280 285  
 Met Cys Glu Met Ile Lys Arg Asn Asp Leu Leu Arg Leu Ser Glu Thr  
 290 295 300  
 Val Ile Lys Pro Phe Tyr Tyr Gln Arg Val Gln Glu Thr Ile Ala Ile  
 305 310 315 320  
 Ile Arg Arg Tyr Leu Pro Glu Asn Arg Cys Leu Ile His Lys Pro Glu  
 325 330 335  
 Gly Ala Ile Phe Leu Trp Leu Trp Phe Lys Asp Leu Pro Ile Thr Thr  
 340 345 350  
 Lys Gln Leu Tyr Gln Arg Leu Lys Ala Arg Gly Val Leu Met Val Pro  
 355 360 365  
 Gly His Asn Phe Phe Pro Gly Leu Asp Lys Pro Trp Pro His Thr His  
 370 375 380  
 Gln Cys Met Arg Met Asn Tyr Val Pro Glu Pro Glu Lys Ile Glu Ala  
 385 390 395 400  
 Gly Val Lys Ile Leu Ala Glu Glu Ile Glu Arg Ala Trp Ala Glu Ser  
 405 410 415

His

- 77 -

<210> 72  
<211> 8803  
<212> DNA  
<213> Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
pAN294 plasmid

&lt;400&gt; 72

tgcgcccgtc cagggcgcggt ccattcgcca ttcaggctgc gcaactgttg ggaagggcga 60  
tcgggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 120  
ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtga 180  
ttgtaatacg actcactata gggcgaattg ggcccgaagt cgcagtctgg atgaaaagcc 240  
gatgaccgct tttcaggctc gtcagcagct ttttctgtct gtatatgaaa aggaattgtt 300  
tttaacgatg tcagaaacgg caggtcacct tgatgtgttg gaggtgaag aagccatcac 360  
gtcatatttg gaaggaaata ccgtatactt taaaacaatg aagaggtgaa atgggtgaaa 420  
catatagcgg gaaaaaggat ttggataacc ggcgcttcag gagggcttgg agaaagaatc 480  
gcatacttat gcgcggtga aggagcccat gtctgtctgt cggctagacg cgaggatcgt 540  
ttgatagaaa tcaaaaggaa aataaccgag gaatggagcg gacagtgtga gatttttctc 600  
ctggatgtcg gccgcctaga ggatatcgcc cgggtccgcg atcagatcgg ctcgattgat 660  
gtactgatta acaatgcagg ctctcggtata ttgaaacgg ttttagactc tacattggat 720  
gacatgaaag cgatgtttga tgtgaatgtc ttcggcctga tcgcctgtac aaaagcgggtg 780  
cttccgcaaa tgcttgagca aaaaaaggga catatcatca atatcgctc tcaagcgggg 840  
aaaatcgcca caccgaagtc tagcctgtat tccgcgacca aacatgccgt gttaggttac 900  
tcaaacgctt tgcggatgga gctttcgga accggcattt atgtgacaac agtcaaccgg 960  
ggcccgaattc agacggactt tttttccatt gctgataaag gcggggacta cgccaaaaat 1020  
gtcggccgct ggatgcttga tcctgatgac gtggcagctc aaattacagc tgcaattttt 1080  
acgaaaaagc gggagatcaa tcttccgctt ttaatgaatg ccggcactaa gctgtatcag 1140  
ctgtttccag ctcttgtaga aaagctggca ggacgcgcgc tcatgaaaaa ataatgatag 1200  
aactgcctgt ggtggagtgg cttgtttctc acggggcagt ttttgatagt ggaagggaga 1260  
gattgttgaa tgtcagttca ttcagaagtc cttcatgctc tgcttaaaga tccgtttatt 1320  
cagaaactga ttgatgcaga gcctgtattc tgggcaaatt caggcaagaa agagggggcca 1380  
ttaccccgtg cagatgagtg ggcaaccgag atagcggaag cggaaaaaag aatgcagcgg 1440

- 78 -

tttgacacctt acattgccga ggtgtttcct gagacgaaag gcgctaaagg aatcatcgag 1500  
tctccgcttt ttgaggtgca gcatatgaag ggaaagctgg aagcggcata tcagcagcca 1560  
tttcccggaa gatggctttt aaagtgcgac catgagcttc cgatttcagg atcgattaaa 1620  
gcgaggggcg ggatttatga agtgttaaag tatgctgaaa atctcgcgct tcaagaagga 1680  
atgcttcagg aaaccgatga ttaccgcac ttacaggaag agcggtttac cgggtttttc 1740  
tcccgcctatt cgattgctgt cggttcgaca ggaaatctag gtttaagcat cggcatcatc 1800  
ggcgcggcac tcgggtttcg cgtgacagtg catatgtccg ccgatgctaa gcagtggaaa 1860  
aaggatctcc tccgccaaaa gggagtcact gttatggagt acgaaacaga ttacagtga 1920  
gcggtgaacg aaggagacg gcaggcggaa caagatccat tctgttattt tattgatgat 1980  
gaacattctc gtcagctgtt cttaggatat gctgttgctg caagccgatt aaaaacacag 2040  
cttgactgta tgaatataaa gccaaagtctt gagacgccct tgtttggtga tctgccgtgc 2100  
ggagtcggcg gaggaccggg cgggtgtagca tttgggctga agcttttata cggagatgat 2160  
gttcatgtgt ttttcgcaga accaactcat tcaccttgta tgctgttagg gctttattca 2220  
ggacttcacg agaagatctc cgtccaggat atcgycctgg ataatacagac ggctgctgac 2280  
ggacttgccg tagggaggcc gtcaggattt gtcgcaagc tgattgaacc gcttctgagc 2340  
ggctgttata cggtagagga caatacgctt tatactttgc ttcatatgct ggctgtatct 2400  
gaagataaat atttagagcc ctctgctctt gctggcatgt tcgggccggg tcagcttttt 2460  
tcgacagaag agggaaggcg ctatgctcag aaatataaga tggaacatgc cgtacatgtc 2520  
gtctggggaa cgggaggaag catggtcca aaagatgaaa tggctgcgta taaccgaatc 2580  
ggtgctgatt tgctaaaaaa acgaaatgga aaataagcag acagtgaaaa ggttttccgt 2640  
tacaatcttt gtaagggttt taacctacag agagtcaggt gtaaacagtg aaaaataaag 2700  
aacttaacct acatacttta tatacacagc acaatcgga gtcttggtct ggttttgggg 2760  
ggcatttgct gattgctgt tctgaagaag aggcaaaagc tgtggaagga ttgaatgatt 2820  
atctatctgt tgaagaagtg gagacgatct atattccgct tgttcgcttg cttcatttac 2880  
atgtcaagtc tgcggtgaa cgcaataagc atgtcaatgt ttttttgaag caccacatt 2940  
cagccaaaat tccgtttatt atcggcattg ccggcagtgt cgcagtcgga aaaagcacga 3000  
cggcgcgat cttgcagaag ctgctttcgc gtttgctga ccgtccaaaa gtgagcctta 3060  
tcacgacaga tggtttttta tttcctactg ccgagctgaa aaagaaaaat atgatgtcaa 3120  
gaaaaggatt tctgaaagc tatgatgtaa aggcgctgct cgaatttttg aatgacttaa 3180

aatcaggaaa ggacagcgta aaggccccgg tgtattccca tctaacctat gaccgcgagg 3240  
aaggtgtggt cgagggtgta gaacaggcgg atattgtgat tattgaaggc attaattgttc 3300  
ttcagtcgcc caccttggag gatgaccggg aaaaccccgcg tatTTTTgtt tccgatttct 3360  
ttgatttttc gatttatgtg gatgcggagg aaagccggat tttcacttgg tatttagagc 3420  
gttttcgcct gcttcgggaa acagcttttc aaaatcctga ttcatttttt cataaattta 3480  
aagacttgct cgatcaggag gctgacgaga tggcagcctc gatttgggag agtgtcaacc 3540  
ggccgaattt atatgaaaat attttgccaa ctaaattcag gtcagatctc attttgcgta 3600  
agggagacgg gcataaggct gaggaagtgt tggtaaggag ggtatgaaat gtgctgcagc 3660  
tcgagcaata gttaccctta ttatcaagat aagaaagaaa aggatttttc gctacgctca 3720  
aatcctttta aaaaacacaa aagaccacat tttttaatgt ggtctttatt cttaactaa 3780  
agcaccatt agttcaacaa acgaaaattg gataaagtgg gatattttta aaatatatat 3840  
ttatgttaca gtaattattga cttttaaaaa aggattgatt ctaatgaaga aagcagacaa 3900  
gtaagcctcc taaattcact ttagataaaa atttaggagg catatcaaat gaactttaat 3960  
aaaattgatt tagacaattg gaagagaaaa gagatattta atcattattt gaaccaacaa 4020  
acgactttta gtataaccac agaaattgat attagtgttt tataccgaaa cataaaacaa 4080  
gaaggatata aattttaccc tgcattttatt ttcttagtga caagggtgat aaactcaaat 4140  
acagctttta gaactggta caatagcgac ggagagttag gttattggga taagttagag 4200  
ccactttata caatttttga tgggtgtatct aaaacattct ctggtatttg gactcctgta 4260  
aagaatgact tcaaagagtt ttatgattta tacctttctg atgtagagaa atataatgg 4320  
tcgggggaaat tgtttcccaa aacacctata cctgaaaatg ctttttctct ttctattatt 4380  
ccatggactt catttactgg gtttaactta aatatcaata ataatagtaa ttaccttcta 4440  
cccattatta cagcaggaaa attcattaat aaaggtaatt caatatattt accgctatct 4500  
ttacaggtac atcattctgt ttgtgatggg tatcatgcag gattgtttat gaactctatt 4560  
caggaattgt cagataggcc taatgactgg cttttataat atgagataat gccgactgta 4620  
ctttttacag tcggttttct aatgtcacta acctgccccg ttagttgaag aaggttttta 4680  
tattacagct gtcgactcgt gatcttcgga caggctgttc agctttttct caatgcgatc 4740  
cagctgcgct tttcggtttt tcgcatactt gaagcctgta acagccgcaa agacgacagc 4800  
ggcaaataata ataaatacaa acagctgaaa catcacatca cctatattca tgttcttcac 4860  
ctcatgtttg cgggagagat tcattctctt ccgtttttta tttaaagcgg cttttccaga 4920

cgggaaacggt gttttgtggt ctccattttc atttgccgat aggcgaacgc taaaaatggc 4980  
aggccgagca gggtaatgcc gctcaggaca gaaaaaatat aaatcggccg gccagcgcca 5040  
aacagggtcta tacatatccc cccgacccaa gggccgatga cgtttccgag ctgtggaaaa 5100  
ccgattgccc cgaaataagt gccttttaat cctgggttttg caatctggtc tacatacaaa 5160  
tccatcatag agaataaaaag cacttcgccg attgtaaatg tgatgacaat catcacaatt 5220  
gatggaaacac cgtgtgatac ggtgaaaatg gccatgctga tgctaaccat cacattaccg 5280  
agcatcagag aacaaagcgg cgaaaaccgt tttgcaaat ggacaatggg aaattgcgtc 5340  
gccaacacaa cgattgcgtt taatgtcagc atcagcccat acagcttcgt tccattgccg 5400  
atcaaggggt tctgcgccat atactgaggg aatgtggaac tgaattgtga gtagccgaag 5460  
gtgcatagcg taatgccgac caaagcaatg gtaaaaagat aatccttttg cgtgaccata 5520  
aacgcttccc gcacgctcat atttcgggac tgggctggtg ctgataagga tggatgtttt 5580  
ttaaattgga gggcaagcac aattccgtat agtccgtaaa tgactgcagg caccaaaaag 5640  
ggcgtagtgc attgcgatga gccgaaatat aggccaagca caggtccgaa gacaacgccg 5700  
atattaatag ccgcatagcg taaattaaaa actagcagtc tcgttttttc ttctgtcata 5760  
tcagacaaca aggcctttga agcgggctca aacagtgatt tgcaaagacc gtttaatgcg 5820  
tttactacaa aaaacaccca gagattagat gctgccgcaa agcctgcaaa taccagcatc 5880  
catccgaaaa tcgatacaag catcatgttt tttctgccga atttatctga gatatatccg 5940  
ccgtaaaagc ttgcgaggat gccgactgat gagctcgcgg cgatgaccag ccctgcatag 6000  
gaagctgatg cgccttggac ggctgtcaaa taaatcgcta aaaaaggaat gctcatcgat 6060  
gttgccattc tgccgaaaat ggttccgatt ataattgtac gcgttgatg catagcttga 6120  
gtattctata gtgtcaccta aatagcttgg cgtaatcatg gtcatagctg tttcctgtgt 6180  
gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 6240  
cctgggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 6300  
tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 6360  
gcggtttgcg tattgggcgc tcttcgctt cctcgctcac tgactcgctg cgctcggtcg 6420  
ttcggctgcg gcgagcggtc tcagctcact caaaggcggg aatacggtta tccacagaat 6480  
caggggataa cgcaggaaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 6540  
aaaaggccgc gttgctggcg tttttcgata ggctccgccc ccctgacgag catcacaaaa 6600  
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 6660

- 81 -

cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt 6720  
cgccttttct cccttcggga agcgtggcgc tttctcatag ctacgctgt aggtatctca 6780  
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagccccg 6840  
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtgaaga cagacttat 6900  
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 6960  
cagagttctt gaagtgggtg cctaactacg gctacactag aaggacagta tttggtatct 7020  
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 7080  
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 7140  
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaaacgaaa 7200  
actcacgtta agggattttg gtcattgagat tatcaaaaag gatcttcacc tagatccttt 7260  
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tgggtctgaca 7320  
gttaccaatg cttaatcagt gaggcacctt tctcagcgat ctgtctatct cgttcattcca 7380  
tagttgcttg actccccgtc gtgtagataa ctacgatacg ggagggtta ccatctggcc 7440  
ccagtgtgc aatgataccg cgagaccac gctcaccggc tccagattta tcagcaataa 7500  
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc 7560  
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 7620  
acgttggttg cattgctaca ggcattcgtg tgtcacgctc gtcgtttggt atggcttcat 7680  
tcagctccgg ttcccaacga tcaaggcgag ttacatgato ccccatgttg tgcaaaaaag 7740  
cggttagctc cttcggctct ccgattcgtt tcagaagtaa gttggccgca gtgttatcac 7800  
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 7860  
ctgtgactgg tgagtactca accaagtcatt tctgagaata ccgcgcccgg cgaccgagtt 7920  
gctcttgccc ggcgtcaata cgggataata gtgtatgaca tagcagaact taaaagtgc 7980  
tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 8040  
ccagttcgat gtaaccact cgtgcacca actgatcttc agcatctttt actttcacca 8100  
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca 8160  
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 8220  
gttattgtct catgagcgga tacatatctg aatgtattta gaaaaataaa caaatagggg 8280  
ttccgcgcac atttccccga aaagtgccac ctgtatgcgg tgtgaaatac cgcacagatg 8340  
cgtaaggaga aaataccgca tcaggcgaaa ttgtaaacgt taatattttg taaaattcg 8400



- 82 -

cggttaaataat ttgttaaatac agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc 8460  
cttataaatac aaaagaatag accgagatag ggttgagtgt tgttccagtt tggaacaaga 8520  
gtccactatt aaagaacgtg gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg 8580  
atggcccact acgtgaacca tcacccaaat caagtttttt gcggtcgagg tgccgtaaag 8640  
ctctaaatcg gaaccctaaa gggagccccc gatttagagc ttgacgggga aagccggcga 8700  
acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg 8760  
tagcggtcac gctgcgcgta accaccacac ccgccgcgct taa 8803

&lt;210&gt; 73

&lt;211&gt; 8320

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
pAN296 plasmid

&lt;400&gt; 73

tgcgcgcgcta cagggcgcgct ccattcgcca ttcaggctgc gcaactgttg ggaagggcga 60  
tcggtgcggg cctcttcgct attacgccag ctggcgaaaag ggggatgtgc tgcaaggcga 120  
ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtga 180  
ttgtaatacg actcactata gggcgaattg ggcccgcgct cgcatgctgg atgaaaagcc 240  
gatgaccgct tttcaggctc gtcagcagct ttttctgct gtatatgaaa aggaattggt 300  
tttaacgatg tcagaaacgg caggtcacct tgatgtgttg gaggtgaag aagccatcac 360  
gtcatatttg gaaggaaata ccgtatactt taaaacaatg aagaggtgaa atgggtgaaa 420  
catatagcgg gaaaaaggat ttggataacc ggcgcttcag gagggcttgg agaaagaatc 480  
gcatacttat gcgcggctga aggagcccat gtctgtctgt cggctagacg cgaggatcgt 540  
ttgatagaaa tcaaaaggaa aataaccgag gaatggagcg gacagtgtga gatttttcct 600  
ctggatgtcg gccgcctaga ggatatcgcc cgggtccgcg atcagatcgg ctcgattgat 660  
gtactgatta acaatgcagg cttcggata tttgaaacgg ttttagactc tacattggat 720  
gacatgaaag cgatgtttga tgtgaatgtc ttcggcctga tcgcctgtac aaaagcggtg 780  
cttccgcaaa tgcttgagca aaaaaaggga catatcatca atatcgctc tcaagcgggg 840  
aaaatcgcca caccgaagtc tagcctgtat tccgcgacca aacatgccgt gttaggttac 900  
tcaaacgctt tgccgatgga gctttcggga accggcattt atgtgacaac agtcaaccgg 960  
ggcccgattc agacggactt tttttccatt gctgataaag gcggggacta cgccaaaaat 1020

gtcggccgct ggatgcttga tcttgatgac gtggcagctc aaattacagc tgcaattttt 1080  
acgaaaaagc gggagatcaa tcttccgcgt ttaatgaatg ccggcactaa gctgtatcag 1140  
ctgtttccag ctctttaga aaagctggca ggacgcgcgc tcatgaaaaa ataatgatag 1200  
aactgcctgt ggtggagtgg cttgtttctc acggggcagt ttttgatagt ggaagggaga 1260  
gattgttgaa tgtcagttca ttcagaagtc cttcatgctc tgcttaaaga tccgtttatt 1320  
cagaaactga ttgatgcaga gcctgtattc tgggcaaatt caggcaagaa agaggggcca 1380  
ttaccccgtg cagatgagtg ggcaaccgag atagcggag cggaaaaaag aatgcagcgg 1440  
tttgacactt acattgccga ggtgtttcct gagacgaaag gcgctaaagg aatcatcgag 1500  
tctccgcttt ttgaggtgca gcatatgaag ggaaagctgg aagcggcata tcagcagcca 1560  
tttcccggaa gatggctttt aaagtgcgac catgagcttc cgatttcagg atcgattaaa 1620  
gcgaggggag ggatttatga agtggttaaag tatgctgaaa atctcgcgct tcaagaagga 1680  
atgcttcagg aaaccgatga ttaccgcata ttacaggaag agcgggttac cgggtttttc 1740  
tcccgtatt cgattgctgt cggttcgaca ggaaatctag gtttaagcat cggcatcctc 1800  
ggcgcggcac tcgggtttcg cgtgacagtg catatgtccg ccgatgctaa gcagtggaaa 1860  
aaggatctcc tccgccaaaa gggagtcact gttatggagt acgaaacaga ttacagtga 1920  
gcggtgaacg aaggagagcg gcaggcggaa caagatccat tctgttattt tattgatgat 1980  
gaacattctc gtcagctggt cttaggatat gctgttgctg caagccgatt aaaaacacag 2040  
cttgactgta tgaatataaa gccaaagtct gagacgcctc tgtttggtga tctgccgtgc 2100  
ggagtcggcg gaggaccggg cgggtgtagca tttgggctga agcttttata cggagatgat 2160  
gttcatgtgt ttttcgcaga accaactcat tcaccttgta tgcgttagg gctttattca 2220  
ggacttcacg agaagatctc cgtccaggat atcggcctgg ataatacagac ggctgctgac 2280  
ggacttgccg tagggaggcc gtcaggattt gtcggcaagc tgattgaacc gcttctgagc 2340  
ggctgttata cggtagagga caatacgctt tatactttgc ttcatatgct ggctgtatct 2400  
gaagataaat atttagagcc ctctgctctt gctggcatgt tcgggccggt tcagcttttt 2460  
tcgacagaag agggaaggcg ctatgctcag aaatataaga tggaacatgc cgtacatgct 2520  
gtctggggaa cgggaggaag catggttcca aaagatgaaa tggctgcgta taaccgaatc 2580  
ggtgctgatt tgctaaaaaa acgaaatgga aaataagcag acagtgaaaa ggttttccgt 2640  
tacaatcttt gtaagggttt taacctacag agagtcaggt gtaaacagtg aaaaataaag 2700  
aacttaacct acatacttta tatacacagc acaatcggga gtcttctgca gctcgagcaa 2760

tagttaccct tattatcaag ataagaaaga aaaggatttt tcgctacgct caaatccttt 2820  
aaaaaacac aaaagaccac attttttaat gtggtcttta ttcttcaact aaagcaccca 2880  
ttagttcaac aaacgaaaat tggataaagt gggatatttt taaaatatat atttatgtta 2940  
cagtaatatt gacttttaaa aaaggattga ttctaataaa gaaagcagac aagtaagcct 3000  
cctaaattca ctttagataa aaatttagga ggcatatcaa atgaacttta ataaaattga 3060  
tttagacaat tggaagagaa aagagatatt taatcattat ttgaaccaac aaacgacttt 3120  
tagtataacc acagaaattg atattagtggt ttataaccga aacataaaac aagaaggata 3180  
taaattttac cctgcattta ttttcttagt gacaagggtg ataaactcaa atacagcttt 3240  
tagaactggt tacaatagcg acggagaggt aggttattgg gataagttag agccacttta 3300  
tacaattttt gatggtgtat ctaaaacatt ctctggtatt tggactcctg taaagaatga 3360  
cttcaaagag ttttatgatt tatacctttc tgatgtagag aaatataatg gttcggggaa 3420  
attgtttccc aaaacaccta tacctgaaaa tgctttttct ctttctatta ttccatggac 3480  
ttcattttact gggtttaact taaatatcaa taataatagt aattaccttc taccattat 3540  
tacagcagga aaattcatta ataaaggtaa ttcaatatat ttaccgctat ctttacaggt 3600  
acatcattct gtttgatgat gttatcatgc aggattgttt atgaactcta ttcaggaatt 3660  
gtcagatagg cctaataact ggcttttata atatgagata atgccgactg tactttttac 3720  
agtcggtttt ctaatgtcac taacctgcc cgtttagttga agaaggtttt tatattacag 3780  
ctgtcgacta aggtcgagga agtggttgga aggagggtat gaaatgtgca tcatattgaa 3840  
ctgtatgtct ctgatttgga ggcgtctagg cggttttggg gctggttctt aaaagaactt 3900  
ggttataaag agtatcaaaa atggagctca ggcacagct ggaagaaaga tcgtttttac 3960  
ctagtgattg tgcaggcgaa agagccattt ctagagccgg aataccatag atgccgagtc 4020  
ggtctgaacc atctcgcat tcatgctgaa tccaagcttc aagtcgatca gatgactgaa 4080  
aaattgacgg caaaaggcta tcgtgtgttg taccgagaca ggcacccctt tgccggagga 4140  
gacgggcatt atgcagtctt ttgtgaggat ccagaccgga ttaaggtaga gctcgttgcc 4200  
ccaagctggt aatcgtgatc ttccggacagg ctgttcagct ttttctcaat gcgatccagc 4260  
tgcgcttttc ggtttttcgc atacttgaag cctgtaacag ccgcaaagac gacagcggca 4320  
aatataataa atacaaacag ctgaaacatc acatcaccta tattcatgtt cttcacctca 4380  
tgtttgcggg agagattcat tctcttcctg tttttattta aagcggcttt tccagacggg 4440  
aacggtgttt tgtggtctcc attttcattt gccgataggc gaacgctaaa aatggcaggc 4500

cgagcagggg aatgccgctc aggacagaaa aaatataaat cggccggcca gcgccaacaa 4560  
ggctctataca tatccccccg acccaagggc cgatgacgtt tccgagctgt ggaaaaccga 4620  
ttgccccgaa ataagtgcct ttaatacctg gttttgcaat ctggtctaca taaaaatcca 4680  
tcatagagaa taaaagcact tcgccgattg taaatgtgat gacaatcacc acaattgatg 4740  
gaacaccgtg tgatacgtg aaaatggcca tgctgatgct aaccatcaca ttaccgagca 4800  
tcagagaaca aagcggcgaa aaccgttttg caaaatggac aatgggaaat tgcgtcgcca 4860  
acacaacgat tgcgtttaat gtcagcatca gcccatacag cttcgttcca ttgccgatca 4920  
aggggttctg cgccatatac tgagggaatg tggaactgaa ttgtgagtag ccgaagggtc 4980  
atagcgtaat gccgaccaa gcaatggtaa aaagataatc cttttgcgtg accataaacg 5040  
cttcccgcac gctcatatct cgggactggg ctggtgctga taaggatgga tgttttttaa 5100  
attggagggc aagcacaatt ccgtatagtc cgtaaataac tgcaggcacc aaaaagggcg 5160  
tagtcgattg cgatgagccg aaatataggc caagcacagg tccgaagaca acgccgatat 5220  
taatagccgc atagcgtaaa ttaaaaacta gcagtctcgt tttttcttct gtcataatcag 5280  
acaacaaggc ctttgaagcg ggctcaaaaa gtgatttgca aagaccgttt aatgcgttta 5340  
ctacaaaaaa caccagaga ttagatgctg ccgcaaagcc tgcaaatacc agcatccatc 5400  
cgaaaatcga tacaagcatc atgttttttc tgccgaattt atctgagata tatccgccgt 5460  
aaaagcttgc gaggatgccg actgatgagc tcgcggcgat gaccagccct gcataggaag 5520  
ctgatgcgcc ttggacggct gtcaataaaa tcgctaaaaa aggaatgctc atcgatgttg 5580  
ccattctgcc gaaaatgggt ccgattataa ttgtaacgcg ttggatgcat agcttgagta 5640  
ttctatagtg tcacctaaat agcttggcgt aatcatggtc atagctgttt cctgtgtgaa 5700  
attgttatcc gtcacaatt ccacacaaca tacgagccgg aagcataaag tgtaaagcct 5760  
ggggtgccta atgagtgagc taactcacat taattgcgtt gcgctcactg cccgctttcc 5820  
agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagagggc 5880  
gtttgcgtat tgggcgctct tccgttccct cgctcaactga ctgcgtgcgc tcggtcgttc 5940  
ggctgcggcg agcggatatc gctcactcaa aggcggtaat acggttatcc acagaatcag 6000  
gggataacgc aggaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa 6060  
aggccgcgtt gctggcgttt ttcgataggc tccgcccccc tgacgagcat caaaaaatc 6120  
gacgctcaag tcagaggtgg cgaaaccgga caggactata aagataccag gcgtttcccc 6180  
ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg 6240

cctttctccc ttcggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt 6300  
cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccgtt cagcccgacc 6360  
gctgcgctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac gaattatcgc 6420  
cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 6480  
agttcttgaa gtggtggcct aactacggct aactagaag gacagtattt ggtatctgcg 6540  
ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa 6600  
ccaccgctgg tagcgggtgt tttttgttt gcaagcagca gattacgcgc agaaaaaag 6660  
gatctcaaga agatccttg atcttttcta cggggtctga cgctcagtgg aacgaaaact 6720  
cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa 6780  
atcaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt 6840  
accaatgctt aatcagttag gcacctatct cagcgatctg tctatttcgt tcatccatag 6900  
ttgctgact ccccgctgtg tagataacta cgatacggga gggcttacca tctggcccca 6960  
gtgctgcaat gataccgcga gaccacgct caccggctcc agatttatca gcaataaacc 7020  
agccagccgg aagggccyag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt 7080  
ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg 7140  
ttgttgcat tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg gcttcattca 7200  
gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg 7260  
ttagctcctt cggctcctcg atcgttgtoa gaagtaagt ggccgcagtg ttatcactca 7320  
tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg 7380  
tgactggtga gtactcaacc aagtcattct gagaataaccg cgccggcgga ccgagttgct 7440  
cttgcccggc gtcaatacgg gataatagt tatgacatag cagaacttta aaagtgtca 7500  
tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca 7560  
gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg 7620  
tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac 7680  
ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt 7740  
attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacia ataggggttc 7800  
cgcgcacatt tccccgaaaa gtgccacctg tatgcggtgt gaaataccgc acagatgcgt 7860  
aaggagaaaa taccgcatca ggcgaaattg taaacgttaa tattttgta aaattcgcgt 7920  
taaataattg ttaaatacgc tcatttttta accaataggc cgaaatcggc aaaatccctt 7980

- 87 -

ataaatcaaa agaatagacc gagatagggt tgagtgttgt tccagtttgg aacaagagtc 8040  
 cactattaaa gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg 8100  
 gccactacg tgaaccatca cccaaatcaa gttttttgctg gtcgaggtgc cgtaaagctc 8160  
 taaatcgga ccctaaaggg agccccgat ttagagcttg acgggggaaag ccggcgaaacg 8220  
 tggcgagaaa ggaagggaag aaagcgaaag gagcgggctg tagggcgctg gcaagtgtag 8280  
 cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa 8320

&lt;210&gt; 74

&lt;211&gt; 250

&lt;212&gt; PRT

&lt;213&gt; Clostridium acetobutylicum

&lt;400&gt; 74

Asn	Lys	Arg	Ala	Ala	Phe	Met	Leu	Leu	Leu	Phe	Leu	Arg	Ser	Val	Leu
1				5					10					15	
Lys	Val	Ile	Leu	Val	Leu	Asp	Val	Gly	Asn	Thr	Asn	Ile	Val	Leu	Gly
			20					25					30		
Ile	Tyr	Asn	Asp	Thr	Lys	Leu	Thr	Ala	Glu	Trp	Arg	Leu	Ser	Thr	Asp
		35					40					45			
Val	Leu	Arg	Ser	Ala	Asp	Glu	Tyr	Gly	Ile	Gln	Val	Met	Asn	Leu	Phe
	50					55					60				
Gln	Gln	Asp	Lys	Leu	Asp	Pro	Thr	Leu	Val	Glu	Gly	Val	Ile	Ile	Ser
65					70					75					80
Ser	Val	Val	Pro	Asn	Ile	Met	Tyr	Ser	Leu	Glu	His	Met	Ile	Arg	Lys
				85					90					95	
Tyr	Phe	Lys	Ile	Asn	Pro	Leu	Val	Val	Gly	Pro	Gly	Ile	Lys	Thr	Gly
			100					105					110		
Ile	Asn	Ile	Lys	Tyr	Asp	Asn	Pro	Lys	Glu	Val	Gly	Ala	Asp	Arg	Ile
	115						120					125			
Val	Asn	Ala	Val	Ala	Ala	His	Glu	Ile	Tyr	Lys	Arg	Ser	Leu	Ile	Ile
	130					135					140				
Ile	Asp	Phe	Gly	Thr	Ala	Thr	Thr	Phe	Cys	Ala	Val	Arg	Glu	Asn	Gly
145					150				155						160
Asp	Tyr	Leu	Gly	Gly	Ala	Ile	Cys	Pro	Gly	Ile	Lys	Val	Ser	Ser	Glu
			165						170					175	
Ala	Leu	Phe	Glu	Lys	Ala	Ala	Lys	Leu	Pro	Arg	Val	Glu	Leu	Ile	Lys
			180					185					190		
Pro	Ala	Tyr	Ala	Ile	Cys	Lys	Asn	Thr	Ile	Ser	Ser	Ile	Gln	Ser	Gly
			195				200						205		

- 88 -

Ile Val Tyr Arg Tyr Leu Arg Gln Val Lys Tyr Leu Phe Glu Lys Leu  
 210 215 220

Lys Glu Asn Leu Pro Asp Gly Arg Arg Thr Arg Thr Ser Leu Val Leu  
 225 230 235 240

Ala Thr Gly Gly Leu Ala Lys Leu Ile Asn  
 245 250

<210> 75

<211> 258

<212> PRT

<213> Rhodobacter capsulatus

<400> 75

Met Leu Leu Cys Ile Asp Cys Gly Asn Thr Asn Thr Val Phe Ser Val  
 1 5 10 15

Trp Asp Gly Thr Asp Phe Ala Ala Thr Trp Arg Ile Ala Thr Asp His  
 20 25 30

Arg Arg Thr Ala Asp Glu Tyr Phe Val Trp Leu Asn Thr Leu Met Gln  
 35 40 45

Leu Lys Gly Leu Gln Gly Arg Ile Ser Glu Ala Ile Ile Ser Ser Thr  
 50 55 60

Ala Pro Arg Val Val Phe Asn Leu Arg Val Leu Cys Asn Arg Tyr Phe  
 65 70 75 80

Asp Cys Arg Pro Tyr Val Val Gly Lys Pro Gly Cys Glu Leu Pro Val  
 85 90 95

Ala Pro Arg Val Asp Pro Gly Thr Thr Val Gly Pro Asp Arg Leu Val  
 100 105 110

Asn Thr Val Ala Gly Tyr Asp Arg His Gly Gly Asp Leu Ile Val Val  
 115 120 125

Asp Phe Gly Thr Ala Thr Thr Phe Asp Val Val Ala Pro Asp Gly Ala  
 130 135 140

Tyr Ile Gly Gly Val Ile Ala Pro Gly Val Asn Leu Ser Leu Glu Ala  
 145 150 155 160

Leu His Met Ala Ala Ala Ala Leu Pro His Val Asp Val Thr Lys Pro  
 165 170 175

Gln Gly Val Ile Gly Thr Asn Thr Val Ala Cys Ile Gln Ser Gly Val  
 180 185 190

Tyr Trp Gly Tyr Ile Gly Leu Val Glu Gly Ile Val Arg Gln Ile Arg  
 195 200 205

Met Glu Arg Asp Arg Pro Met Lys Val Ile Ala Thr Gly Gly Leu Ala  
 210 215 220

- 89 -

Ser Leu Phe Asp Leu Gly Phe Asp Leu Phe Asp Lys Val Glu Asp Asp  
 225 230 235 240

Leu Thr Met His Gly Leu Arg Leu Ile Phe Asp Tyr Asn Lys Gly Leu  
 245 250 255

Gly Ala

<210> 76

<211> 10801

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Recombinant  
 pAN240 plasmid

<400> 76

gaattttgcg gccgcttcga aagctgtaat ataaaaacct tcttcaacta acgggggcagg 60  
 ttagtgacat tagaaaaccg actgtaaaaa gtacagtcgg cattatctca tattataaaa 120  
 gccagtcatt aggcctatct gacaattcct gaatagagtt cataaacaat cctgcatgat 180  
 aaccatcaca aacagaatga tgtacctgta aagatagcgg taaatatatt gaattacctt 240  
 tattaatgaa ttttcttgct gtaataatgg gtagaaggta attactatta ttattgatat 300  
 ttaagttaaa ccagtaaat gaagtcctatg gaataataga aagagaaaaa gcattttcag 360  
 gtataggtgt tttgggaaac aatttccccg aaccattata tttctctaca tcagaaaggt 420  
 ataaatcata aaactctttg aagtcattct ttacaggagt ccaaatacca gagaatgttt 480  
 tagatacacc atcaaaaatt gtataaagtg gctctaactt atcccaataa cctaactctc 540  
 cgtcgctatt gtaaccagtt ctaaaagctg tatttgagtt tatcaccctt gtcactaaga 600  
 aaataaatgc agggtaaaat ttatatcctt cttgttttat gtttcggtat aaaacactaa 660  
 tatcaatttc tgtggttata ctaaaagtcg tttgttggtt caaataatga ttaaatatct 720  
 cttttctctt ccaattgtct aaatcaattt tattaaagtt catttgatat gcctcctaaa 780  
 tttttatcta aagtgaattt aggaggctta cttgtctgct ttcttcatta gaatcaatcc 840  
 ttttttaaaa gtcaatatta ctgtaacata aatatatatt ttaaaaatat ccactttat 900  
 ccaattttcg tttgttgaac taatgggtgc tttagttgaa gaataaagac cacattaaaa 960  
 aatgtggtct tttgtgtttt tttaaaggat ttgagcgtag cgaaaaatcc ttttctttct 1020  
 tatcttgata ataagggtaa ctattgaatt cggtaccaag agtttgtaga aacgcaaaaa 1080  
 ggccatccgt caggatggcc ttctgcttaa tttgatgcct ggcagtttat ggcgggcgtc 1140



ctgcccgccca ccctccgggc cgttgcttcg caacgttcaa atccgctccc ggcggatttg 1200  
tcctactcag gagagcgttc accgacaaac aacagataaa acgaaaggcc cagtcttttcg 1260  
actgagcctt tcgttttatt tgatgcctgg cagttcccta ctctcgcagtg gggagacccc 1320  
acactaccat cggcgctacg gcgtttcact tctgagttcg gcatgggggc aggtggggacc 1380  
accgcgctac tgccgccagg caaattctgt tttatcagac cgcttctgcg ttctgattta 1440  
atctgtatca ggctgaaaat cttctctcat ccgccaaaac aggatcctac ggaaatggag 1500  
cggcaaaacc gttttactct caaaatctta aaagaaaacc cccgataaag ggggcttttc 1560  
ttctacaaaa ttgtacgggc tggttcgttc ccagcattt gttcaatttt gttttgatca 1620  
ttcagaacag ccactttcgg ctcatggctt gccgcttctt gatcagacat cattttgtag 1680  
gaaataataa tgaccttatt tccttcctgc acaaggcgtg cggctgcacc gtttaagcat 1740  
atgacgccgc ttccccgttt accaggaata atatacgttt caagacgtgc tccattatta 1800  
ttattcaciaa tttgtacttt ttcattagga agcattccca cagcatcaat gagatcttca 1860  
tcaattgtaa tgcttccac atagttcagg tttgcttccg taacagttgc cctgtgaagt 1920  
ttgccgctca tcattgttcg atacatatta tattctctcc atttctcgaa tatcaataat 1980  
gatattatct attaaacgcg cttttgaaaa agcaactgca acagcgagaa tcacttttcc 2040  
agcaatttca ttcacaggct cgagttccgg ataggaataa agctctacat agtctatggg 2100  
tccgctagtc gtttcaatga tatcttttgc agcttttatt accgcttcag gatctctttc 2160  
accggcttgg acaagttccg cacttgtttg aagggcccca tacagcttag gcgcttcttt 2220  
tctttctca gctgttaagt atacattgcg agagcttttg gctaagccgt cttcctctct 2280  
gaccgtatcg acaggaacca attcaatatt catgaagaag tcgctgatta acccatcaac 2340  
aacagctacc tgctgcgcatt cttttaaac gaaataggca cgagtcggct tgactagatt 2400  
gaaaagcttc gtcagtacga tcgcgacccc gtcaaaatgt ctttctcttg agcgcccgca 2460  
taacacgtct gtgcgtcttt ctacatgaat cgtgacattc ttttcaccgg gatacatatt 2520  
atgagcatct ggcgtaaaaa gaatatcgac tccggcgttt tctgcaagag ctgcatcccg 2580  
ctcaatatcg cgcggatatg cttcaaaatc ttcattaggg ccgaattgtg caggattcac 2640  
aaaaatactc ataataacgg cgtcgttttc ttgtcttgct ttgtctgcta aggttaaatg 2700  
cccctcatgc agaaacccca tcgtcggaac aaatccgatt gacttgccct ctgaatggta 2760  
ttgttttatg gcttctttca gctgtgaaat atcagtaatc tgtctcatct tattttcccc 2820  
cgtacaagcc gtcaagcact gtctggttca tttgaaagga atgcttttgt tcagggaaaag 2880

cacgatgtct tacatcctga acatatccgc tgattgctgt ttcgatgggt tcatcaatgc 2940  
gcgtatattg ctttacaaat ttaggtgttc tctcaacacc gtggccgata atatcatgat 3000  
aaacgagaac ttgtccgtcc gctttcacac cagccccgat tccaatgacc ggtatgctta 3060  
gcgtctcggc aatttttggt gtgagttctg ccggcacaca ttcagcaca agcatcatag 3120  
ctcctgcttc ttgcatttt atactgtctt ctattaattt tttggcgctt tgttcgtctt 3180  
tgccctgtac tttatagccg ccagtagcg cgactgactg cgggtgtcaa cctaagtgc 3240  
tgactactgg aatgcctcca agcgtcaatg cgcgaatgga ttcaaacacg ccttctccgc 3300  
cctcaagctt cagtgcgtca gctccgcttt cctgaacgat agccgctgca tttttcagcg 3360  
tatcttcctt agacaggtga taagacataa acggcatatc tgtcacaata aaggatttcg 3420  
gcgcacccct tttaacggct tttgtatgat ggatcatgtc cgcaactgtc acaccgacag 3480  
ttgaatcaag gccgaggacg accattccaa gtgaatcacc gactaaaatc atgtcaactc 3540  
ccgcttgctc agcaagttta gctgccggat aatcataagc ggtcagcatg acaatcggtt 3600  
cttcagactc cttcattttt agaaaatcca gttttgtttt catgttttct cctccttcta 3660  
gagcgtcctg ctgttggtta gattattata ccacaccttg tagataaagt caacaacttt 3720  
ttgcaaaatt tttcaggaat tttagcagag gttgttctgg atgtagaaca aaacatcttt 3780  
ccgctcttgt gctgttagga tatctttctt ggaagctagg taggcctcga gttatggcag 3840  
ttggttaaaa ggaaacaaaa agaccgtttt cacacaaaac ggtctttttc gatttctttt 3900  
tacagtcaca gccacttttg caaaaaccgg acagcttcat gccttataac tgctgtttcg 3960  
gtcgacgatg atctgccgtt ttcttctgca agccaaaaaa cttccgtta caacgagaag 4020  
gattcttcac tttctaaagt tcggcgagtt tcatccctct gtcccagtc ttttttggat 4080  
caaggcagac tgctgcaatg tctatctatt ttaataatag gtgcagttcg caggcgatac 4140  
tgcccaatgg aagtatacca aaatcaacgg gcttgtagca acacattagc ccaattcgat 4200  
atcggcagaa tagatttttt taatgccttc gttcgtttct aaaagcagaa cgccttcac 4260  
atctatacct aacgccttac cgtaaaaggt tccgtttaac gttctggctc tcatattagt 4320  
gcccaatacg agcgcatagc tttcccataa agcttaatc ggcgtaaatc cgtgcgtcat 4380  
ataatcccgg taccgtttct caaagcatag taaaatatgc tggatgacgc cggcccgatc 4440  
aattttttcc ccagcagctt ggctgaggct tgctcgcatg tccttcaatt catctggaaa 4500  
atcattaggc tgctggttaa cgttaatgcc gatcccaatg atcactgaac gtacgcggtc 4560  
ttcttcagcc tgcatttcg ttaggatacc gactgttttt tttccgttaa tcaaaatata 4620

atttggccat ttaatatccg tttggatgcc tgctgcctct tctattccct gcacaacagc 4680  
tactgcagca agcagagtca gctgcggtgt tttttggagc ggaatgtcag gccgcaaaat 4740  
caggctcatc caaacaccgt ttcttctctg agaatgccat accctagaca ttcgggccct 4800  
tccggctggt tgtttgtcag ccaccacaag ggtgccttcc ggtgcgttat tattcgcgag 4860  
ctcatgagcc gttttttgcg tgcttgaaag aacgtcatgg taaataagat gctggcccat 4920  
cacttccgtt tttaatcaa aacgaatttc gctttcactg agttttccgg gttttttgat 4980  
gagccgatat ccttttcttc taacggcttc tacttcataa ccctctttcc gaagctcttc 5040  
aatatgcttc cacacagcag ttcttgaaca gccgagagca tcaactgattt tttggccgga 5100  
aataaattca ttgccggcct gagaaaataa ttcaataagg tcttttctta atgttgaccg 5160  
catgtcttca gccactctc tatgtgtttc ttttgattgg agagcttccc tgtcacaaca 5220  
gcttgctcga tccactgtaa ttcttctgac acccattttc cggccggccg gtttcgaagc 5280  
gcaagcaagt ccttaccgt gatataaga tccttaaggc ttttgatcgg caggttttga 5340  
taagcgtact gaatgtcctt cagtttcttt tcatccagtt ttcgttttg ccgaagctgc 5400  
gatattttgg ccgctgagag cagtgtttt ttcccagctc tgtacattgt cattgcgtca 5460  
aggctctggc caaacgtatc ggcaatgtga atggcttctc tgatcacttt tcccgggagc 5520  
ttccaggctt tcaggaaaag gggcgcgctc ttcaaaacta tgccaagggt aattaaaga 5580  
gcagcccaaa gctcctcacg ggatgttaaa gagaagaatg gaaactcact cgttgaaatc 5640  
aggttttctc gtttatgata aaaaccagga agctcttcat acaatctcgt ttgaatgagt 5700  
gtttgaagcg cctggcgaga agctcttccc tgcagcaatt tctcaaactc tatagttttt 5760  
cgttcgactg aaacatggga gaggagtgat ttttctttcg caatggcttc ttctgtttcc 5820  
ggtgaaagcg taaagccaag ctggctcata aagcgtacgg ctctcagcat acgaagcgca 5880  
tcctcttgaa atctatctc aggttttcca acggttcgaa tcactttctg atcaatatct 5940  
ttcttgccgc caaaataatc aagcaccttc ccgtccgctg tcatggccat cgcattgatc 6000  
gttaaattctc tgcgttttag atctcttct aatgatgaga taaattgcac ttctgacggt 6060  
cttctgaaat caacataatc agattcagtc cggaatgtcg tgacttcata ggtttcatcc 6120  
tcccagagca caataatggt cccgtgctct ttgcctacat caacagtccg ctgaaacagc 6180  
cgttctactt gatcaggtgc cgcactgtc gcgatatcga catctccgat cgttcgtttc 6240  
atatagctgt cacgaactgc gccccgaca aaataagcct gatggccgc ttcgattaag 6300  
atgcggagca cgggaagtgc tttgataaaa actttttcca tgtgatcact ccggttctgc 6360

taaatcggca taaatctggt catactggct gacaattttt ttagaagaaa attcattttc 6420  
aagcatctct attgccgcct ttgtaaaacg attgcttagc tgttcacott ctaaaatgct 6480  
catcgcgcg gctggtgagg ccgtaacatc accgacatcc accaaaaatc cgctcacatt 6540  
gttttttata acctcaggga taccgccaat gtttgttcca atacaaggca ctccgcaagc 6600  
catcgcttca agcaggacaa ggccaaagct ttctttttca gatagcagca gcttcaaate 6660  
gctaatagaa taaagatctt caacacggc ttgattttcca agcattaaga cttggtcttc 6720  
caagccatat ttcttgataa gctcgcaggc tgctgatttc tccggaccgt ctccgactaa 6780  
aagcagcttc gctttcggtt tgccagcgat attgcggaac acacggatga catcctgcac 6840  
gcgtgcagcc actggttaaca ggattagcag agcgaggatg gtaggcgggtg ctacagagtt 6900  
cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggtg tctgcgctct 6960  
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 7020  
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 7080  
tcaagaagat cctttgatct ttctacggg gtctgacgt cagtggaaacg aaaactcacg 7140  
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaatta 7200  
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 7260  
atgcttaate agtgaggcac ctatctcagc gatctgtcta ttctgttcac ccatagttgc 7320  
ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc 7380  
tgcaatgata ccgcgagacc cagctcacc ggctccagat ttatcagcaa taaaccagcc 7440  
agccggaagg gccgagcgca gaagtgggtc tgcaacttta tccgcctcca tccagtctat 7500  
taattgttgc cgggaagcta gagtaagtag ttccgcagtt aatagtttgc gcaacgttgt 7560  
tgccattgct acaggcatcg tgggtgtcacg ctctcggtt ggtatggctt cattcagctc 7620  
cggttcccaa cgatcaaggc gagttacatg atccccatg ttgtgcaaaa aagcggttag 7680  
ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt 7740  
tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac 7800  
tggtgagtag tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg 7860  
cccggcgta atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat 7920  
tggaanaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc 7980  
gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc 8040  
tggtgagca aaaacaggaa ggcaaatgc cgcaaaaag ggaataaggc cgacacggaa 8100

atgttgaata ctcataactct tcctttttca atattattga agcatttatc agggttattg 8160  
tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg 8220  
cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac 8280  
ctataaaaaat aggcgtatca cgaggccctt tcgtctcgca tgcggatcag tgagggtttg 8340  
caactgcggg tcaaggatct ggatttcgat cacggcaega tcatcgtgcg ggagggcaag 8400  
ggctccaagg atcgggcctt gatgttacct gagagcttgg caccagcct gcgcgagcag 8460  
gggaattgat ccggtggatg accttttgaa tgacctttaa tagattatat tactaattaa 8520  
ttggggaccc tagaggcccc cttttttatt ttaaaaattt tttcacaaaa cggtttacaa 8580  
gcataacggg ttttgctgcc cgcaaacggg ctgttctggt gttgctagtt tgttatcaga 8640  
atcgcagatc cggcttcagg tttgccggct gaaagcgcta tttcttcag aattgccatg 8700  
atTTTTTccc cacgggaggc gtcactggct cccgtgttgt cggcagcttt gattcgataa 8760  
gcagcatcgc ctgtttcagg ctgtctatgt gtgactgttg agctgtaaca agttgtctca 8820  
ggtgttcaat ttcattgtct agttgctttg ttttactggt ttcacctgtt ctattaggtg 8880  
ttacatgctg ttcattctgt acattgtcga tctgttcatt gtgaacagct ttaaattgcac 8940  
caaaaactcg taaaagctct gatgtatcta tcttttttac accgttttca tctgtgcata 9000  
tggaacagttt tccctttgat atctaacggg gaacagttgt tctacttttg tttgttagtc 9060  
ttgatgcttc actgatagat acaagagcca taagaacctc agatccttcc gtatttagcc 9120  
agtatgttct ctagtgtggt tcgttgtttt tgcgtgagcc atgagaacga accattgaga 9180  
tcatgcttac tttgcatgct actcaaaaat tttgcctcaa aactgggtgag ctgaattttt 9240  
gcagttaaag catcgtgtag tgtttttctt agtccgttac gtaggtagga atctgatgta 9300  
atggttgttg gtattttgtc accattcatt tttatctggt tgttctcaag ttcggttacg 9360  
agatccattt gtctatctag ttcaacttgg aaaatcaacg tatcagtcgg gcggcctcgc 9420  
ttatcaacca ccaatttcat attgctgtaa gtgtttaaat ctttacttat tggtttcaaa 9480  
accattgggt taagcctttt aaactcatgg tagttatttt caagcattaa catgaactta 9540  
aattcatcaa ggctaattct tatatttgcc ttgtgagttt tcttttgtgt tagttctttt 9600  
aataaccact cataaatcct catagagtat ttgttttcaa aagacttaac atgttccaga 9660  
ttatatttta tgaatttttt taactggaaa agataaggca atatctcttc actaaaaact 9720  
aattctaatt tttcgcttga gaacttggca tagtttgtcc actggaaaat ctcaaagcct 9780  
ttaaccaaag gattcctgat ttccacagtt ctcgatcatca gctctctggt tgcttttagct 9840

- 95 -

aatacaccat aagcattttc cctactgatg ttcattcatct gagcgtattg gttataagtg 9900  
 aacgataccg tccgttcttt ccttgtaggg ttttcaatcg tgggggtgag tagtgccaca 9960  
 cagcataaaa ttagcttggt ttcattgctcc gtttaagtcatt agcgactaat cgctagttca 10020  
 tttgctttga aaacaactaa ttcagacata catctcaatt ggtctagggtg attttaatca 10080  
 ctataccaat tgagatgggc tagtcaatga taattactag tccttttctt ttgagttgtg 10140  
 ggtatctgta aattctgcta gacctttgct ggaaaacttg taaattctgc tagaccctct 10200  
 gtaaattccg ctagaccttt gtgtgttttt tttgtttata ttcaagtggg tataatttat 10260  
 agaataaaga aagaataaaa aaagataaaa agaataagatc ccagccctgt gtataactca 10320  
 ctactttagt cagttccgca gtattacaaa aggatgtcgc aaacgctgtt tgctcctcta 10380  
 caaaacagac cttaaaaccc taaaggctta agtagcacc cgcgaagctc gggcaaactc 10440  
 ctgaatatcc cttttgtctc cgaccatcag gcacctgagt cgctgtcttt ttcgtgacat 10500  
 tcagttcgct gcgctcacgg ctctggcagt gaatgggggt aaatggcact acaggcgctt 10560  
 tttatggatt catgcaagga aactacccat aatacaagaa aagcccgta cgggcttctc 10620  
 agggcggttt atggcgggtc tgctatgtgg tgctatctga cttttcgctg ttcagcagtt 10680  
 cctgccctct gattttccag tctgaccact tcggattatc ccgtgacagg tcattcagac 10740  
 tggctaattgc acccagtaag gcagcgggtat catcaacagg cttaccgctc ttactgtcaa 10800  
 c 10801

&lt;210&gt; 77

&lt;211&gt; 8654

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
 pAN236 plasmid

&lt;400&gt; 77

ctcgaggcct acctagcttc caagaaagat atcctaacag cacaagagcy gaaagatggt 60  
 ttgttctaca tccagaacaa cctctgctaa aattcctgaa aaattttgca aaaagttggt 120  
 gactttatct acaagggtgtg gtataataat cttaacaaca gcaggacgct ctagaggagg 180  
 agacatcatg aaaattggaa ttatcggcgg aggtccggtt ggtcttttat gcgcctatta 240  
 tttgtcactt tatcacgacg tgactgttgt gacgaggcgg caagaacagg ctgcggccat 300  
 tcagtctgaa ggaatccggc ttataaaagg cggggaggaa ttcagggtctg attgcagtgc 360

- 96 -

ggacacgagt atcaattcgg actttgacct gcttgctgtg acagtgaagc agcatcagct 420  
tcaatctgtt ttttcgtcgc ttgaacgaat cgggaagacg aatatattat ttttgcaaaa 480  
cggcatgggg catatccacg acctaaaaga ctggcacgtt ggccattcca tttatgttgg 540  
aatcgttgag cacggagctg taagaaaatc ggatacagct gttgatcata caggcctagg 600  
tgcgataaaa tggagcgcgt tcgacgatgc tgaaccagac cggctgaaca tcttgtttca 660  
gcataaccat tcggattttc cgatttatta tgagacggat tggtagcgtc tgctgacggg 720  
caagctgatt gtaaagcgt gtattaatcc ttttaactgcg ttattgcaag tgaaaaatgg 780  
agaactgctg acaacgccag cttatctggc ttttatgaag ctggtatttc aggaggcatg 840  
ccgcatttta aaacttgaaa atgaagaaa ggcttgggag cgggttcagg ccgtttgtgg 900  
gcaaacgaaa gagaatcgtt catcaatgct ggttgacgtc attggaggcc ggacagcgga 960  
agctgacgcc attatcggat acttattgaa ggaagcaagt cttcaaggtc ttgatgccgt 1020  
ccacctagag tttttatatg gcagcatcaa agcattggag cgaaatacaa acaaagtctt 1080  
ttgagctttt tcggtaacat gctatactca tttcggatca ctaactattt attggaggat 1140  
cctgttttgg cggatgagag aagattttca gcctgataca gattaatca gaacgcagaa 1200  
gcggtctgat aaaacagaat ttgcctggcg gcagtagcgc ggtggtccca cctgacccca 1260  
tgccgaactc agaagtgaia cgcgtagcg ccgatggtag tgtggggtct ccccatgcga 1320  
gagtagggaa ctgccaggca tcaaataaaa cgaaaggctc agtcgaaaga ctgggccttt 1380  
cgttttatct gttgtttgtc ggtgaacgct ctctgagta ggacaaatcc gccgggagcg 1440  
gatttgaacg ttgcgaagca acggcccgga gggtagcggg caggacgccc gccataaact 1500  
gccaggcatc aaattaagca gaaggccatc ctgacggatg gcctttttgc gtttctacaa 1560  
actcttggtc cccagaaaaa gcggcaaaa cggctgttaa aaaagcgaaa tcgaagaagc 1620  
tgtctgccgc taagacggaa tatcaaaagc gttctgctgt tgtgtcatct ttaaaagtca 1680  
cagccgatga atcccagcaa gatgtcctaa aatacttgaa caccagaaa gataaaggaa 1740  
atgcagacca aattcattct tattatgtgg tgaacgggat tgctgttcat gcctcaaaag 1800  
aggttatgga aaaagtgggt cagtttcccg aagtggaaaa ggtgcttctt aatgagaaac 1860  
ggcagctttt taagtcattc tccccattta atatgaaaaa agcacagaaa gctattaaag 1920  
caactgacgg tgtggaatgg aatgtagacc aaatcgatgc cccaaaagct tgggcacttg 1980  
gatatgatgg aactggcacg gttgttgcgt ccattgatac cggggtggaa tggaatcatc 2040  
cggcattaaa agagaaatat cgcggatata atccggaaaa tcctaatgag cctgaaaatg 2100

aaatgaactg gtatgatgcc gtagcaggcg aggcaagccc ttatgatgat ttggctcatg 2160  
gaacccacgt gacaggcacg atggtgggct ctgaacctga tggaaacaaat caaatcgggtg 2220  
tagcacctgg cgcaaaatgg attgctgtta aagcgttctc tgaagatggc ggcaactgatg 2280  
ctgacattttt ggaagctggt gaatgggttt tagcaccaaa ggacgcggaa ggaaatcccc 2340  
acccggaat ggctcctgat gttgtcaata actcatgggg agggggctct ggacttgatg 2400  
aatggtacag agacatggtc aatgcctggc gttcggccga tattttccct gagttttcag 2460  
cggggaatac ggatctcttt attcccggcg ggccctggttc tatcgcaaat ccggcaaaact 2520  
atccagaatc gtttgcaact ggagcgactg agaattccaa ttccccatgg agagaaaaga 2580  
aaatcgctaa tgttgattac tttgaacttc tgcataattct tgaatttaaa aaggctgaaa 2640  
gagtaaaaga ttgtgctgaa atattagagt ataaacaaaa tcgtgaaaca ggcgaaagaa 2700  
agttgtatcg agtggtgttt tgtaaatcca ggctttgtcc aatgtgcaac tggaggagag 2760  
caatgaaaca tggcattcag tcacaaaagg ttgttgctga agttattaaa caaaagccaa 2820  
cagttcgttg gttgtttctc acattaacag ttaaaaatgt ttatgatggc gaagaattaa 2880  
ataagagttt gtcagatag gctcaaggat ttcgccgaat gatgcaatat aaaaaatta 2940  
ataaaaatct tgttggtttt atgcgtgcaa cggaagtgc aataaataat aaagataatt 3000  
cttataatca gcacatgcat gtattggtat gtgtggaacc aacttatttt aagaatacag 3060  
aaaactacgt gaatcaaaaa caatggattc aattttggaa aaaggcaatg aaattagact 3120  
atgatccaaa tgtaaaagtt caaatgattc gaccgaaaaa taaatataaa tcggatatac 3180  
aatcggaat tgacgaaact gcaaaatatc ctgtaaagga tacggatttt atgaccgatg 3240  
atgaagaaaa gaatttgaaa cgtttgtctg atttgaggga aggtttacac cgtaaaaggt 3300  
taatctccta tgggtggttg ttaaaagaaa tacataaaaa attaaacctt gatgacacag 3360  
aagaaggcga tttgattcat acagatgatg acgaaaaagc cgatgaagat ggatttttcta 3420  
ttattgcaat gtggaattgg gaacggaaaa attattttat taaagagtag ttcaacaaac 3480  
gggcatatt gttgtataag tgatgaaata ctgaatttaa aacttagttt atatgtggta 3540  
aaatgtttta atcaagttta ggaggaatta attatgaagt gtaatgaatg taacagggtt 3600  
caattaaaag agggaagcgt atcattaacc ctataaacta cgtctgccct cattattgga 3660  
gggtgaaatg tgaatacatc ctattcacia tcgaatttac gacacaacca aattttaatt 3720  
tggttttgca ttttatcttt ttttagcgta ttaaatgaaa tggttttgaa cgtctcatta 3780  
cctgatattg caaatgattt taataaacca cctgcgagta caaactgggt gaacacagcc 3840



- 98 -

tttatgttaa ccttttccat tggaacagct gtatatggaa agctatctga tcaattaggc 3900  
atcaaaaggt tactcctatt tggaattata ataaattggt tcgggtcggg aattgggttt 3960  
gttggccatt ctttcttttc cttacttatt atggctcgtt ttattcaagg ggctggtgca 4020  
gctgcatttc cagcactcgt aatggttgta gttgcgcgct atattccaaa ggaaaatagg 4080  
ggtaaagcat ttggtcttat tggatcgata gtagccatgg gagaaggagt cgggccagcg 4140  
attggtggaa tgatagccca ttatattcat tggtcctatc ttctactcat tccatgata 4200  
acaattatca ctgttcggtt tcttatgaaa ttattaaaga aagaagtaag gataaaaggt 4260  
cattttgata tcaaaggaat tatactaata tctgtaggca ttgtattttt tatgttggtt 4320  
acaacatcat atagcatttc ttttcttctc gttagcgtgc tgcattcct gatatttgta 4380  
aaacatatca ggaaagtaac agatcctttt gttgatcccg gattaggga aaatatacct 4440  
tttatgattg gagttcttg tgggggaatt atatttgaa cagtagcagg gtttgtctct 4500  
atggttcctt atatgatgaa agatgttcac cagctaagta ctgccgaaat cggaagtgt 4560  
attattttcc ctggaacaat gagtgcatt attttcggct acattggtgg gatacttgtt 4620  
gatagaagag gtcctttata cgtgttaaac atcggagtta catttctttc tgttagcttt 4680  
ttaactgctt cctttctttt agaaacaaca tcatggttca tgacaattat aatcgtattt 4740  
gttttaggtg ggctttcgtt caccaaaaca gttatatcaa caattgtttc aagtagcttg 4800  
aaacagcagg aagctggtgc tggaatgagt ttgcttaact ttaccagctt tttatcagag 4860  
ggaacaggta ttgcaattgt aggtggttta ttatccatac cttacttga tcaaagggtg 4920  
ttacctatgg aagttgatca gtcaacttat ctgtatagta atttgttatt acttttttca 4980  
ggaatcattg tcattagttg gctggttacc ttgaatgtat ataaacattc tcaaagggat 5040  
ttctaaatcg ttaagggatc aactttgga gagagttcaa aattgatcct ttttttataa 5100  
cagttcgaag cgccgcaat tcttgaagac gaaagggcct cgtgatacgc ctatttttat 5160  
aggttaatgt catgataata atggtttctt agacgtcagg tggcactttt cggggaaatg 5220  
tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat ccgctcatga 5280  
gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg agtattcaac 5340  
atttccgtgt cgcccttatt cccttttttg cggcattttg cttcctgtt tttgctcacc 5400  
cagaaacgct ggtgaaagta aaagatgctg aagatcagtt ggggtgcacga gtgggttaca 5460  
tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc 5520  
caatgatgag cacttttaaa gttctgctat gtggcgcggt attatccgt attgacgccg 5580

- 99 -

ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt gagtactcac 5640  
cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc agtgctgccca 5700  
taaccatgag tgataaact gcgccaact tacttctgac aacgatcgga ggaccgaagg 5760  
agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat cgttggggaac 5820  
cggagctgaa tgaagccata ccaaagcagc agcgtgacac cacgatgcct gcagcaatgg 5880  
caacaacgtt gcgcaacta ttaactggcg aactacttac tctagcttcc cggcaacaat 5940  
taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg gcccttccgg 6000  
ctggctggtt tattgctgat aaatctggag ccggtgagcg tgggtctcgc ggtatcattg 6060  
cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg acggggagtc 6120  
aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca ctgattaagc 6180  
attggttaact gtcagaccaa gtttactcat atatacttta gattgattta aaacttcatt 6240  
tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc aaaatccctt 6300  
aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa ggatcttctt 6360  
gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca ccgtaccag 6420  
cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta actggcttca 6480  
gcagagcgca gataccaaat actgtccttc tagtgtagcc gtagttaggc caccacttca 6540  
agaactctgt agcaccgcct acatacctcg ctctgctaata cctyttacca gtggctgctg 6600  
ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta ccggataagg 6660  
cgcagcggtc gggctgaacg ggggggtcgt gcacacagcc cagcttgagc cgaacgacct 6720  
acaccgaact gagataccta cagcgtgagc tatgagaaaag cgccacgctt cccgaaggga 6780  
gaaaggcgga caggatatccg gtaagcggca gggtcggaac aggagagcgc acgagggagc 6840  
ttccaggggg aaacgcctgg tatctttata gtccgtgctg gtttcgcccac ctctgacttg 6900  
agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac gccagcaacg 6960  
cggccttttt acggttcctg gccttttgct ggcccttttg tcacatgttc tttcctgcgt 7020  
tatccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat accgctcgcc 7080  
gcagccgaac gaccgagcgc agcagtcag tgagcgagga agcggaagag cgccctgatgc 7140  
ggtattttct ccttacgcat ctgtgcggta tttcacaccg catatgggtgc actctcagta 7200  
caatctgctc tgatgccgca tagttaagcc agtatacact ccgctatcgc tacgtgactg 7260  
ggtcatggct gcgccccgac acccgccaac acccgctgac gcgccctgac gggcttgtct 7320

- 100 -

gctcccgga tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag 7380  
 gttttcaccg tcatcaccga aacgcgcgag gcagctgcgg taaagctcat cagcgtgggtc 7440  
 gtgaagcgat tcacagatgt ctgcctgttc atccgcgtcc agctcgttga gtttctccag 7500  
 aagcgtaaat gtctggcttc tgataaagcg ggccatgtta agggcggttt tttcctgttt 7560  
 ggtaacttga tgcctccgtg taagggggaa tttctgttca tgggggtaat gataccgatg 7620  
 aaacgagaga ggatgctcac gatacgggtt actgatgatg aacatgcccg gttactggaa 7680  
 cgttgtgagg gtaaacaact ggcggtatgg atgcggcggg accagagaaa aatcactcag 7740  
 ggtcaatgcc agcgttctgt taatacagat gtagggtgtc cacagggtag ccagcagcat 7800  
 cctgcgatgc agatccggaa cataatggtg cagggcgctg acttccgcgt ttccagactt 7860  
 tacgaaacac ggaaaccgaa gaccattcat gttgttgctc aggtcgcaga cgttttgcag 7920  
 cagcagtcgc ttcacgttcg ctgcggtatc ggtgattcat tctgctaacc agtaaggcaa 7980  
 ccccgccagc ctagccgggt cctcaacgac aggagcacga tcatgcgcac ccgtggccag 8040  
 gacccaacgc tgcccagat gcgccgcgtg cggctgctgg agatggcgga cgcgatggat 8100  
 atgttctgcc aagggttggg ttgcgcattc acagttctcc gcaagaattg attggctcca 8160  
 attcttgag tggatgaatc gttagcgagg tgccgcggc ttccattcag gtcgaggtgg 8220  
 cccggtcca tgcaccgca cgcaacgcg ggaggcagac aaggatatagg gcggcgcccta 8280  
 caatccatgc caaccgttc catgtgctcg ccgaggcggc ataaatcgcc gtgacgatca 8340  
 gcggtccagt gatcgaagt aggctggtaa gagccgcgag cgatccttga agctgtccct 8400  
 gatggtcgtc atctacctgc ctggacagca tggcctgcaa cgcgggcac ccgatgccgc 8460  
 cggaagcgag aagaatcata atggggaagg ccattccagcc tcgcgtcggc ggccgcttcg 8520  
 tcgaccgaaa cagcagttat aaggcatgaa gctgtccggt ttttgcaaaa gtggctgtga 8580  
 ctgtaaaaag aaatcgaaaa agaccgtttt gtgtgaaaac ggtctttttg tttcctttta 8640  
 accaactgcc ataa 8654

&lt;210&gt; 78

&lt;211&gt; 8093

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
 pAN423 plasmid

&lt;400&gt; 78

ggcggccgct tcgtcgaccg aaacagcagt tataaggcat gaagctgtcc ggtttttgca 60

aaagtggctg tgactgtaaa aagaaatcga aaaagaccgt tttgtgtgaa aacgggtcttt 120  
ttgtttcctt ttaaccaact gccataactc gaggcctacc tagcttccaa gaaagatatc 180  
ctaacagcac aagagcggaa agatgttttg ttctacatcc agaacaacct ctgctaaaat 240  
tcctgaaaaa ttttgcaaaa agttgttgac tttatctaca aggtgtggta taataatctt 300  
aacaacagca ggacgctcta gaaaaggagg aatttaaattg tatcgtagca tgatgagcgg 360  
caaacttcac agggcaactg ttacggaagc aaacctgaac tatgtgggaa gcattacaat 420  
tgatgaagat ctcatgatg ctgtgggaat gcttcctaata gaaaaagtac aaattgtgaa 480  
taataataat ggagcacgtc ttgaaacgta tattattcct ggtaaaccggg gaagcggcgt 540  
catatgctta aacgggtgcag ccgcacgcct tgtgcaggaa ggagataagg tcattattat 600  
ttcctacaaa atgatgtctg atcaagaagc ggcaagccat gagccgaaag tggtgttct 660  
gaatgatcaa aacaaaattg aacaaatgct ggggaacgaa ccagcccgtta caattttgta 720  
aaggatcctg ttttgcgga tgagagaaga ttttcagcct gatacagatt aaatcagaac 780  
gcagaagcgg tctgataaaa cagaatttgc ctggcggcag tagcgcggtg gtcccacctg 840  
accccatgcc gaactcagaa gtgaaacgcc gtacgcgcca tggtagtggtg ggggtctcccc 900  
atgcgagagt agggaaactgc caggcatcaa ataaaacgaa aggctcagtc gaaagactgg 960  
gcctttcggtt ttatctgttg tttgtcgggtg aacgctctcc tgagtaggac aaatccgccg 1020  
ggagcggatt tgaacgttgc gaagcaacgg cccggagggt ggcgggcagg acgcccgcga 1080  
taaactgcca ggcatcaaat taagcagaag gccatcctga cggatggcct ttttgcggtt 1140  
ctacaaactc ttggtaccca gaaaaagcgg caaaagcggc tgttaaaaaa gcgaaatcga 1200  
agaagctgtc tgccgctaag acggaatatc aaaagcgttc tgctgttggtg tcatctttaa 1260  
aagtcacagc cgatgaatcc cagcaagatg tcctaaaata cttgaacacc cagaaagata 1320  
aaggaaatgc agaccaaatt cattcttatt atgtggtgaa cgggattgct gttcatgcct 1380  
caaaagaggt tatggaaaaa gtggtgcagt ttcccgaagt ggaaaagggtg cttcctaagt 1440  
agaaacggca gctttttaag tcctcctccc catttaatat gaaaaaagca cagaaagcta 1500  
ttaaagcaac tgacggtgtg gaatggaatg tagaccaaact cgatgccccca aaagcttggg 1560  
cacttgata tgatggaact ggcacggttg ttgctccat tgataccggg gtggaatgga 1620  
atcatccggc attaaaagag aaatatcgcg gatataatcc ggaaaatcct aatgagcctg 1680  
aaaatgaaat gaactggtat gatgcgtag caggcgaggc aagcccttat gatgatttgg 1740  
ctcatggaac ccacgtgaca ggcacgatgg tgggctctga acctgatgga acaaatcaaa 1800

tcggtgtagc acctggcgca aaatggattg ctgttaaagc gttctctgaa gatggcggca 1860  
ctgatgctga cattttggaa gctgggtgaat gggtttttagc accaaaggac gcggaaggaa 1920  
atccccaccc ggaaatggct cctgatgttg tcaataactc atggggaggg ggctctggac 1980  
ttgatgaatg gtacagagac atgggtcaatg cctggcggtc ggccgatatt ttccttgagt 2040  
tttcagcggg gaatacggat ctctttatct ccggcgggcc tggttctatc gcaaataccg 2100  
caaactatcc agaatcggtt gcaactggag cgactgagaa ttccaattcc ccatggagag 2160  
aaaagaaaat cgctaattgt gattactttg aacttctgca tattcttgaa tttaaaaagg 2220  
ctgaaagagt aaaagattgt gctgaaatat tagagtataa acaaaatcgt gaaacaggcg 2280  
aaagaaagt gtatcgagt tggttttgta aatccaggct ttgtccaatg tgcaactgga 2340  
ggagagcaat gaaacatggc attcagtcac aaaagggtgt tgctgaagtt attaaacaaa 2400  
agccaacagt tcgttggttg tttctcacat taacagttaa aaatgtttat gatggcgaag 2460  
aattaataa gagtttgta gatatggctc aaggatttcg ccgaatgatg caatataaaa 2520  
aaattaataa aaatcttggt ggttttatgc gtgcaacgga agtgacaata aataataaag 2580  
ataattctta taatcagcac atgcatgtat tggatgtgt ggaaccaact tattttaaga 2640  
atacagaaaa ctacgtgaat caaaaacaat ggattcaatt ttggaaaaag gcaatgaaat 2700  
tagactatga tccaaatgta aaagttcaaa tgattcgacc gaaaaataaa tataaatcgg 2760  
atatacaatc ggcaattgac gaaactgcaa aatatcctgt aaaggatacg gattttatga 2820  
ccgatgatga agaaaagaat ttgaaacgtt tgtctgattt ggaggaaggt ttacaccgta 2880  
aaaggttaat ctctatggt ggtttgttaa aagaaatata taaaaatta aacctgatg 2940  
acacagaaga aggcgatttg attcatacag atgatgacga aaaagccgat gaagatggat 3000  
tttctattat tgcaatgtgg aattgggaac ggaaaaatta tttattaaa gagtagttca 3060  
acaaacgggc catattgttg tataagtgat gaaatactga atttaaaact tagtttatat 3120  
gtggtaaaat gttttaatca agtttaggag gaattaatta tgaagtgtaa tgaatgtaac 3180  
agggttcaat taaaagaggg aagcgtatca ttaaccctat aaactacgtc tgccctcatt 3240  
attggagggt gaaatgtgaa tacatcctat tcacaatcga atttacgaca caaccaaatt 3300  
ttaatttggc ttgcatctt atcttttttt agcgtattaa atgaaatggt tttgaacgtc 3360  
tcattacctg atattgcaaa tgattttaat aaaccacctg cgagtacaaa ctgggtgaac 3420  
acagccttta tgtaacctt ttccattgga acagctgtat atggaaagct atctgatcaa 3480  
ttaggcataca aaaggttact cctatttgga attataataa attgtttcgg gtcggtaatt 3540

gggtttgttg gccattcttt cttttcctta cttattatgg ctcgttttat tcaaggggct 3600  
ggtgcagctg catttccagc actcgtaatg gttgtagttg cgcgctatat tccaaaggaa 3660  
aataggggta aagcatttgg tcttattgga tcgatagtag ccatgggaga aggagtcggt 3720  
ccagcgattg gtggaatgat agcccattat attcattggt cctatcttct actcattcct 3780  
atgataacaa ttatcactgt tccgtttctt atgaaattat taaagaaaga agtaaggata 3840  
aaaggtcatt ttgatatcaa aggaattata ctaatgtctg taggcattgt attttttatg 3900  
ttgtttacaa catcatatag catttctttt cttatcgta gcgtgctgac attcctgata 3960  
tttgtaaaac atatcaggaa agtaacagat ccttttgttg atcccggatt agggaaaaat 4020  
atacctttta tgattggagt tctttgtggg ggaattatat ttggaacagt agcagggttt 4080  
gtctctatgg ttccttatat gatgaaagat gttcaccagc taagtactgc cgaaatcgga 4140  
agtgtaatga ttttccttgg aacaatgagt gtcattattt tcggctacat tgggtgggata 4200  
cttgttgata gaagaggtcc tttatacgtg ttaaaccatcg gagttacatt tctttctggt 4260  
agctttttta ctgcttcctt tcttttagaa acaacatcat ggttcatgac aattataatc 4320  
gtatttgttt taggtgggct ttcgttcacc aaaacagtta tatcaacaat tgtttcaagt 4380  
agcttgaaac agcaggaagc tgggtgctgga atgagtttgc ttaactttac cagcttttta 4440  
tcagagggaa caggtattgc aattgtaggt ggtttattat ccataccctt acttgatcaa 4500  
aggttgttac ctatggaagt tgatcagtca acttatctgt atagtaattt gttattacct 4560  
ttttcaggaa tcattgtcat tagttggctg gttacctga atgtatataa acattctcaa 4620  
agggatttct aaatcgtaa gggatcaact ttgggagaga gttcaaaatt gatccttttt 4680  
ttataacagt tcgaagcggc cgcaattctt gaagacgaaa gggcctcgtg atacgcctat 4740  
ttttataggt taatgtcatg ataataatgg tttcttagac gtcagggtggc acttttcggg 4800  
gaaatgtgcg cggaaccctt atttgtttat ttttctaaat acattcaaat atgtatccgc 4860  
tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggag agtatgagta 4920  
ttcaacattt ccgtgtcgcc cttattccct ttttgcggc attttgcctt cctgtttttg 4980  
ctcaccagaa aacgctggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg 5040  
gttacatcga actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac 5100  
gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta tcccgtattg 5160  
acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac ttgggtgagt 5220  
actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg 5280

ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac 5340  
cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt 5400  
gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgcag 5460  
caatggcaac aacgttgcg aaactattaa ctggcgaact acttactcta gcttcccggc 5520  
aacaattaat agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc 5580  
ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta 5640  
tcattgcagc actggggcca gatggtaagc cctcccgat cgtagttatc tacacgacgg 5700  
ggagtccagg aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga 5760  
ttaagcattg gtaactgtca gaccaagttt actcatatat actttagatt gatttaaaac 5820  
ttcattttta atttaaaagg atctaggtga agatcctttt tgataatctc atgacaaaaa 5880  
tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat 5940  
cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc 6000  
taccagcggg gggttggttg ccggatcaag agctaccaac tctttttccg aaggtaactg 6060  
gcttcagcag agcgcagata ccaaatactg tccttctagt gtagccgtag ttaggccacc 6120  
acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaatectg ttaccagtgg 6180  
ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg 6240  
ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa 6300  
cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg 6360  
aaggagaaaa ggcggacagg tatccggtaa gcggcagggg cggaacagga gagcgcacga 6420  
gggagcttcc agggggaaac gcctggtatc tttatagtc tgtcgggttt cgccacctct 6480  
gacttgagcg tcgatttttg tgatgctcgt caggggggcg gaggcctatgg aaaaacgcca 6540  
gcaacgcggc ctttttacgg ttcttgccct tttgctggcc ttttgctcac atgttctttc 6600  
ctgcgttate cctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg 6660  
ctcgccgcag ccgaacgacc gagcgcagcg agtcagttag cgaggaagcg gaagagcgcc 6720  
tgatgcggta ttttctcctt acgcatctgt gcggtatttc acaccgcata tgggtgactc 6780  
tcagtacaat ctgctctgat gccgcatagt taagccagta tacactccgc tatcgctacg 6840  
tgactgggtc atggctgcgc cccgacaccc gccaacaccc gctgacgcgc cctgacgggc 6900  
ttgtctgctc ccggcatccg cttacagaca agctgtgacc gtctccggga gctgcatgtg 6960  
tcagaggttt tcaccgtcat caccgaaacg cgcgaggcag ctgcggtaaa gctcatcagc 7020

- 105 -

gtgggtcgtga agcgattcac agatgtctgc ctgttcatcc gcgtccagct cggttgagttt 7080  
 ctccagaagc gttaatgtct ggcttctgat aaagcggggc atgttaaggg cggttttttc 7140  
 ctgtttggtc acttgatgcc tccgtgtaag ggggaatttc tgttcatggg ggtaatgata 7200  
 ccgatgaaac gagagaggat gctcacgata cgggttactg atgatgaaca tgcccgggta 7260  
 ctggaacgtt gtgagggtaa acaactggcg gtatggatgc ggcgggacca gagaaaaatc 7320  
 aetcagggtc aatgccagcg ctctgttaac acagatgtag gtgttccaca gggtagccag 7380  
 cagcatcctg cgatgcagat ccggaacata atgggtgcagg gcgctgactt ccgcgtttcc 7440  
 agactttacg aaacacggaa accgaagacc attcatgttg ttgctcaggt cgcagacgtt 7500  
 ttgcagcagc agtcgcttca cgttcgctcg cgtatcggtg attcattctg ctaaccagta 7560  
 aggcaacccc gccagcctag ccgggtcttc aacgacagga gcacgatcat gcgcacccgt 7620  
 ggccaggacc caacgctgcc cgagatgcgc cgcgtgcggc tgctggagat ggccggacgcg 7680  
 atggatatgt tctgccaagg gttggtttgc gcattcacag ttctccgcaa gaattgattg 7740  
 gctccaattc ttggagtggg gaatccgtta gcgagggtgcc gccggcttcc attcagggtcg 7800  
 aggtggcccc gctccatgca ccgcgacgca acgcggggag gcagacaagg tatagggcgg 7860  
 cgcctacaat ccatgccaac ccgttccatg tgctcgccga ggccggcataa atcgccgtga 7920  
 cgatcagcgg tccagtgate gaagttaggc tggttaagagc cgcgagcgat ccttgaagct 7980  
 gtccctgatg gtcgtcatct acctgcctgg acagcatggc ctgcaacgcg ggcatcccga 8040  
 tgccgcccga agcgagaaga atcataatgg ggaaggccat ccagcctcgc gtc 8093

&lt;210&gt; 79

&lt;211&gt; 8098

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
 pAN429 plasmid

&lt;400&gt; 79

ggcggccgct tcgtcgaccg aaacagcagt tataaggcat gaagctgtcc ggtttttgca 60  
 aaagtggctg tgactgtaaa aagaaatcga aaaagaccgt tttgtgtgaa aacgggtcttt 120  
 ttgtttcctt ttaaccaact gccataactc gaggcctacc tagcttccaa gaaagatatt 180  
 ctaacagcac aagagcggaa agatgttttg ttctacatcc agaacaacct ctgctaaaaat 240  
 tcctgaaaaa ttttgcaaaa agttgttgac tttatctaca aggtgtggta taataatctt 300



aacaacagca ggaagctcta gattagaaag gaggtttaat taatgtatcg tacgatgatg 360  
agcggcaaac ttcaacagggc aactgttacg gaagcaaac cc tgaactatgt gggaagcatt 420  
acaattgatg aagatctcat tgatgctgtg ggaatgcttc ctaatgaaaa agtacaaatt 480  
gtgaataata ataatggagc acgtcttgaa acgtatatta ttcttggtaa acggggaagc 540  
ggcgtcatat gcttaaaccg tgcagccgca cgccttggtc aggaaggaga taaggtcatt 600  
attatttctt acaaaatgat gtctgatcaa gaagcggcaa gccatgagcc gaaagtggct 660  
gttctgaatg atcaaaacaa aattgaacaa atgctgggga acgaaccagc ccgtacaatt 720  
ttgtaaagga tcctgttttg gcggatgaga gaagattttc agcctgatac agattaaatc 780  
agaacgcaga agcgggtctga taaaacagaa ttgacctggc ggagtagcg cgggtggccc 840  
acctgacccc atgccgaact cagaagtga acgcctgagc gccgatggta gtgtggggtc 900  
tccccatgag agagtaggga actgccaggc atcaaataaa acgaaaggct cagtcgaaag 960  
actgggcctt tcgttttatc tgttgtttgt cggatgaacg tctctgagt aggacaaatc 1020  
cgccgggagc ggatttgaac gttgcgaagc aacggcccg aggtggcg gcaggacgcc 1080  
cgccataaac tgccaggcat caaattaagc agaaggccat cctgacggat ggcttttttg 1140  
cgtttctaca aactcttggg acccagaaaa agcggcaaaa gcggctgtta aaaaagcgaa 1200  
atcgaagaag ctgtctgccg ctaagacgga atatcaaaag cgttctgctg ttgtgtcatc 1260  
tttaaaagtc acagccgatg aatcccagca agatgtccta aaatacttga acaccagaa 1320  
agataaagga aatgcagacc aaattcattc ttattatgtg gtgaacggga ttgctgttca 1380  
tgctcaaaa gaggttatgg aaaaagtggg gcagtttccc gaagtggaaa aggtgcttcc 1440  
taatgagaaa cggcagcttt ttaagtcac cccccattt aatatgaaaa aagcacagaa 1500  
agctattaaa gcaactgagc gtgtggaatg gaatgtagac caaatcgatg ccccaaaagc 1560  
ttgggcactt ggatatgatg gaactggcac gggtgttgcg tccattgata ccgggggtgga 1620  
atggaatcat ccggcattaa aagagaaata tcgcggatat aatccggaaa atcctaata 1680  
gcctgaaaat gaaatgaact ggtatgatgc cgtagcaggc gaggcaagcc cttatgatga 1740  
tttggctcat ggaaccacg tgacaggcac gatggtgggc tctgaacctg atggaacaaa 1800  
tcaaatcggg ttagcacctg gcgcaaaatg gattgctgtt aaagcgttct ctgaagatgg 1860  
cggcactgat gctgacattt tggaagctgg tgaatgggtt ttagcaccaa aggacgcgga 1920  
aggaaatccc caccggaaa tggctcctga tgttgtcaat aactcatggg gagggggctc 1980  
tggaactgat gaatgtgaca gagacatggg caatgcctgg cgttcggccg atattttccc 2040

tgagttttca gcggggaata cggatctctt tattcccggc gggcctgggt ctatcgcaaa 2100  
tccggcaaac tatccagaat cgtttgcaac tggagcgact gagaattcca attccccatg 2160  
gagagaaaag aaaatcgcta atgttgatta ctttgaactt ctgcatattc ttgaatttaa 2220  
aaaggctgaa agagtaaaag attgtgctga aatattagag tataaacaaa atcgtgaaac 2280  
aggcgaaaga aagttgtatc gagtgtgggt ttgtaaatcc aggctttgtc caatgtgcaa 2340  
ctggaggaga gcaatgaaac atggcattca gtcacaaaag gttgttgctg aagttattaa 2400  
acaaaagcca acagttcgtt ggttgtttct cacattaaca gttaaaaatg tttatgatgg 2460  
cgaagaatta aataagagtt tgtcagatat ggctcaagga tttcgccgaa tgatgcaata 2520  
taaaaaaatt aataaaaatc ttgttggttt tatgctgca acggaagtga caataaataa 2580  
taaagataat tcttataatc agcacatgca tgtattggta tgtgtggaac caacttattt 2640  
taagaataca gaaaactacg tgaatcaaaa acaatggatt caattttgga aaaaggcaat 2700  
gaaattagac tatgatccaa atgtaaaagt tcaaatgatt cgaccgaaaa ataaatataa 2760  
atcggatata caatcggcaa ttgacgaaac tgcaaaatat cctgtaaagg atacggattt 2820  
tatgaccgat gatgaagaaa agaatttgaa acgtttgtct gatttgaggg aaggtttaca 2880  
ccgtaaaagg ttaatctcct atgggtggtt gttaaaagaa atacataaaa aattaaacct 2940  
tgatgacaca gaagaaggcg atttgattca tacagatgat gacgaaaaag ccgatgaaga 3000  
tggattttct attattgcaa tgtggaattg ggaacggaaa aattatttta ttaaagagta 3060  
gttcaacaaa cgggccatat tgttgataaa gtgatgaaat actgaattta aaacttagtt 3120  
tatatgtggt aaaatgtttt aatcaagttt aggaggaatt aattatgaag tgtaatgaat 3180  
gtaacagggt tcaattaaaa gagggaagcg tatcattaac cctataaact acgtctgccc 3240  
tcattattgg aggtgaaat gtgaatacat cctattcaca atcgaattta cgacacaacc 3300  
aaattttaat ttggctttgc attttatctt tttttagcgt attaaatgaa atggttttga 3360  
acgtctcatt acctgatatt gcaaatgatt ttaataaacc acctgcgagt acaaactggg 3420  
tgaacacagc ctttatgtta accttttcca ttggaacagc tgtatatgga aagctatctg 3480  
atcaattagg catcaaaagg ttactcctat ttggaattat aataaattgt ttcgggtcgg 3540  
taattgggtt tgttgccat tctttctttt ccttacttat tatggctcgt tttattcaag 3600  
gggctgggtc agctgcattt ccagcactcg taatggttgt agttgcgcgc tatattccaa 3660  
aggaaaatag gggtaaagca tttggtctta ttggatcgat agtagccatg ggagaaggag 3720  
tcgggtccagc gattggtgga atgatagccc attatattca ttggtcctat cttctactca 3780

ttcctatgat aacaattatc actgttccgt ttcttatgaa attattaaag aaagaagtaa 3840  
ggataaaagg tcattttgat atcaaaggaa ttatacta at gtctgtaggc attgtatttt 3900  
ttatgttggt tacaacatca tatagcattt cttttcttat cgttagcgtg ctgtcattcc 3960  
tgatatttgt aaaacatc aggaaagtaa cagatcctt tgttgatccc ggattaggga 4020  
aaaatatacc ttttatgatt ggagttctt gtgggggaat tatatttgga acagtagcag 4080  
ggtttgtctc tatggttctc tatatgatga aagatgttca ccagctaagt actgccgaaa 4140  
tcggaagtgt aattattttc cctggaacaa tgagtgtcat tattttcggc tacattggtg 4200  
ggatacttgt tgatagaaga ggtcctttat acgtgttaaa catcgaggtt acatttctt 4260  
ctgtagctt tttactgct tcctttctt tagaaacaac atcatgggtc atgacaatta 4320  
taatcgtatt tgttttaggt gggctttcgt tcacaaaac agttatatca acaattgtt 4380  
caagtagctt gaaacagcag gaagctgggt ctggaatgag tttgcttaac tttaccagct 4440  
ttttatcaga gggaacaggt attgcaattg taggtggtt attatccata ccctacttg 4500  
atcaaagggt gttacctatg gaagttgatc agtcaactta tctgtatagt aatttggtt 4560  
tacttttttc aggaatcatt gtcattagtt ggtggttac cttgaatgta tataaacatt 4620  
ctcaaaggga tttctaaatc gttaagggt caactttggg agagagttca aaattgatcc 4680  
tttttttata acagttcgaa gcggccgcaa ttcttgaaga cgaaagggcc tcgtgatacg 4740  
cctattttta taggttaatg tcatgataat aatggtttct tagacgtcag gtggcactt 4800  
tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 4860  
tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 4920  
gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttctgt 4980  
ttttgctcac ccagaaacgc tgggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 5040  
agtgggttac atcgaactgg atctcaacag cgtaagatc cttgagagtt ttcgccccga 5100  
agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg 5160  
tattgacgcc gggcaagagc aactcgggtc ccgcatacac tattctcaga atgacttggt 5220  
tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 5280  
cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg 5340  
aggaccgaag gagctaaccg cttttttgca caacatggg gatcatgtaa ctcgccttga 5400  
tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 5460  
tgcagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc 5520

ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc 5580  
ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg 5640  
cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 5700  
gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc 5760  
actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 5820  
aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 5880  
caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 5940  
aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc 6000  
accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 6060  
aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 6120  
ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tctgttacc 6180  
agtggctgct gcagtgggc ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 6240  
accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc ccagcttgga 6300  
gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 6360  
tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg 6420  
cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 6480  
cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa 6540  
cgccagcaac gcggcctttt tacggttcct ggcccttttg tgcccttttg ctcacatgtt 6600  
ctttcctgcg ttatccctg attctgtgga taaccgtatt accgcctttg agtgagctga 6660  
taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga 6720  
gcgcctgatg cggtattttc tccttacgca tctgtgcggc atttcacacc gcatatggtg 6780  
cactctcagt acaatctgct ctgatgccgc atagttaagc cagtatacac tccgctatcg 6840  
ctacgtgact gggtcatggc tgcgccccga caccgcgcaa caccgctga cgcgccctga 6900  
cgggcttgtc tgctcccggc atccgcttac agacaagctg tgaccgtctc cgggagctgc 6960  
atgtgtcaga ggttttcacc gtcatacccg aaacgcgcga ggcagctgcg gtaaagctca 7020  
tcagcgtggc cgtgaagcga ttcacagatg tctgcctgtt catccgcgtc cagctcgttg 7080  
agtttctcca gaagcgtaa tgtctggctt ctgataaagc gggccatgtt aaggcggtt 7140  
ttttcctgtt tggtcacttg atgcctccgt gtaaggggga atttctgttc atgggggtaa 7200  
tgataccgat gaaacgagag aggatgctca cgatacgggt tactgatgat gaacatgccc 7260

- 110 -

ggttactgga acgttgtgag ggtaaacaac tggcgggtatg gatgcggcgg gaccagagaa 7320  
 aaatcactca ggggtcaatgc cagcgcttcg ttaatacaga tgtaggtgtt ccacagggta 7380  
 gccagcagca tcctgcgatg cagatccgga acataatggt gcagggcgct gacttccgcg 7440  
 tttccagact ttacgaaaca cggaaccga agaccattca tgttgttgct caggtcgcag 7500  
 acgttttgca gcagcagtcg cttcacgttc gctcgcgtat cggtgattca ttctgctaac 7560  
 cagtaaggca accccgccag cctagccggg tcctcaacga caggagcacg atcatgcgca 7620  
 cccgtggcca ggaccaacg ctgcccaga tgcgccgct gcggctgctg gagatggcgg 7680  
 acgcgatgga tatgttctgc caagggttg tttgcgcatt cacagttctc cgcaagaatt 7740  
 gattggctcc aattcttga gtggtgaatc cgttagcgag gtgccgcgg cttccattca 7800  
 ggtcgaggtg gcccggtcc atgcaccgcg acgcaacgcg gggaggcaga caaggtatag 7860  
 ggcggcgctt acaatccatg ccaaccgtt ccatgtgctc gccgaggcgg cataaatcgc 7920  
 cgtgacgatc agcgggtccag tgatcgaagt taggctggta agagccgcga gcgatccttg 7980  
 aagctgtccc tgatggtcgt catctacctg cctggacagc atggcctgca acgcgggcat 8040  
 cccgatgccg ccggaagcga gaagaatcat aatggggaag gccatccagc ctgcgctc 8098

&lt;210&gt; 80

&lt;211&gt; 4450

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
 pAN443 plasmid

&lt;400&gt; 80

ccatcgaatg gccagatgat taattcctaa tttttgttga cactctatca ttgatagagt 60  
 tattttacca ctccctatca gtgatagaga aaagtgaat gaatagttcg acaaaaatct 120  
 agattagaaa ggaggattta aatatgagac agattactga tatttcacag ctgaaagaag 180  
 ccataaaaca ataccattca gagggcaagt caatcggatt tgttccgacg atggggtttc 240  
 tgcatgaggg gcatttaacc ttagcagaca aagcaagaca agaaaacgac gccgttatta 300  
 tgagtatttt tgtgaatcct gcacaattcg gccctaataga agattttgaa gcatatccgc 360  
 gcgatattga gcgggatgca gctcttgacg aaaacgccgg agtcgatatt ctttttacgc 420  
 cagatgctca tgatatgtat cccggtgaaa agaatgtcac gattcatgta gaaagacgca 480  
 cagacgtgtt atgcgggcgc tcaagagaag gacattttga cggggtcgcg atcgactga 540  
 cgaagctttt caatctagtc aagccgactc gtgcctattt cggtttaaaa gatgcgcagc 600

- 111 -

aggtagctgt tgttgatggg ttaatcagcg acttcttcat ggatattgaa ttggttcctg 660  
tcgatacggg cagagaggaa gacggcttag ccaaaagctc tcgcaatgta tacttaacag 720  
ctgaggaaag aaaagaagcg cctaagctgt atcgggccct tcaaacaagt gcggaacttg 780  
tccaagccgg tgaaagagat cctgaagcgg tgataaaagc tgcaaaagat atcattgaaa 840  
cgactagcgg aaccatagac tatgtagagc tttattccta tccggaactc gagcctgtga 900  
atgaaattgc tggaaagatg attctcgtg ttgcagttgc tttttcaaaa gcgcgtttaa 960  
tagataatat cattattgat attcgtagaa aggaggtgaa ttaatatgta tcgtacgatg 1020  
atgagcggca aacttcacag ggcaactgtt acggaagcaa acctgaacta tgtgggaagc 1080  
attacaattg atgaagatct cattgatgct gtgggaatgc ttcctaatga aaaagtacaa 1140  
attgtgaata ataataatgg agcacgtctt gaaacgtata ttattcctgg taaacgggga 1200  
agcggcgta tatgcttaaa cgggtgcagcc gcacgccttg tgcaggaagg agataaggta 1260  
attattatct cctacaaaat gatgtctgat caagaagcgg caagccatga gccgaaagt 1320  
gctgttctga atgatcaaaa caaaattgaa caaatgctgg ggaacgaacc agcccgtaca 1380  
attttgtaaa ggatcccccg gggatccctc gaggtcgacc tgcaggggga ccatggtctc 1440  
agcgtttgga gccacccgca gttcgaaaaa taataagctt gacctgtgaa gtgaaaaatg 1500  
gcgcacattg tgcgacattt tttttgtctg ccgtttaccg ctactgcgtc acggatctcc 1560  
acgcgccttg tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc agcgtgaccg 1620  
ctacacttgc cagcgcctta gcgccgctc ctttcgcttt cttcccttc tttctcgcca 1680  
cgttcgccgg ctttccccgt caagctctaa atcgggggct ccctttaggg ttccgattta 1740  
gtgctttacg gcacctcgac ccaaaaaaac ttgattaggg tgatggttca cgtagtgggc 1800  
catcgccctg atagacggtt tttcgccctt tgacgttgga gtccacgttc tttaatagt 1860  
gactcttggt ccaaactgga acaacactca accctatctc ggtctattct tttgatttat 1920  
aagggatttt gccgatttcg gcctattggt taaaaaatga gctgatttaa caaaaattta 1980  
acgcgaattt taacaaaata ttaacgctta caatttcagg tggcactttt cggggaaatg 2040  
tgcgcggaac ccctatttgt ttatttttct aaatacatc aaatatgtat ccgctcatga 2100  
gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg agtattcaac 2160  
atttcgtgt cgcccttatt cccttttttg cggcattttg ctttcctgtt tttgtcacc 2220  
cagaaacgct ggtgaaagta aaagatgctg aagatcagtt ggggtgcacga gtgggttaca 2280  
tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc 2340

- 112 -

caatgatgag cacttttaaa gttctgctat gtggcgcggt attatccgt attgacgccg 2400  
ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt gagtactcac 2460  
cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc agtgctgcc 2520  
taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga ggaccgaagg 2580  
agctaaccgc ttttttgac aacatggggg atcatgtaac tcgccttgat cgttggaac 2640  
cggagctgaa tgaagccata ccaaacgacg agcgtgacac cagcatgcct gtagcaatgg 2700  
caacaacgtt gcgcaaaact ttaactggcg aactacttac tctagcttcc cggcaacaat 2760  
tgatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg gcccttccgg 2820  
ctggctggtt tattgctgat aaatctggag ccggtgagcg tggctctcgc ggtatcattg 2880  
cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg acggggagtc 2940  
aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca ctgattaagc 3000  
attggtagga attaatgatg tctcgtttag ataaaagtaa agtgattaac agcgcattag 3060  
agctgcttaa tgaggtcgga atcgaaggtt taacaacccg taaactcgcc cagaagctag 3120  
gtgtagagca gcctacattg tattggcatg taaaaataa gcgggctttg ctcgacgcct 3180  
tagccattga gatgttagat aggcaccata ctacttttg ccctttagaa ggggaaagct 3240  
ggcaagattt ttacgtaat aacgctaaaa gttttagatg tgctttacta agtcatcgcg 3300  
atggagcaaa agtacattta ggtacacggc ctacagaaaa acagtatgaa actctcgaaa 3360  
atcaattagc ctttttatgc caacaagggt ttctactaga gaatgcatta tatgcaactca 3420  
gcgcagtggg gcattttact ttaggttgcg tattggaaga tcaagagcat caagtcgcta 3480  
aagaagaaag ggaaacacct actactgata gtatgccgcc attattacga caagctatcg 3540  
aattatttga tcaccaaggt gcagagccag cttctttatt cggccttgaa ttgatcatat 3600  
gcggattaga aaaacaactt aaatgtgaaa gtgggtctta aaagcagcat aacctttttc 3660  
cgtgatggta acttcactag tttaaaagga tctaggtgaa gatccttttt gataatctca 3720  
tgacaaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga 3780  
tcaaaggatc ttcttgagat cttttttttc tgcgcgtaat ctgctgcttg caaacaaaaa 3840  
aaccaccgct accagcgggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga 3900  
aggtaactgg cttcagcaga gcgcagatac caaatactgt ctttctagt tagccgtagt 3960  
taggccacca cttcaagaac tctgtagcac cgctacata cctcgctctg ctaatcctgt 4020  
taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttgac tcaagacgat 4080

- 113 -

agttaccgga taaggcgag cggtcgggct gaacggggggg ttcgtgcaca cagcccagct 4140  
 tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca 4200  
 cgcttcccga agggagaaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag 4260  
 agcgcacgag ggagcttcca gggggaaaacg cctgggtatct ttatagtcct gtcggggtttc 4320  
 gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggcgg agcctatgga 4380  
 aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca 4440  
 tgacccgaca 4450

&lt;210&gt; 81

&lt;211&gt; 10212

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
 pAN251 plasmid

&lt;400&gt; 81

gcggcgcgcta aaaagagctt gaggatttgc ggagtgaaaa tcagacattg cggaatcagc 60  
 tagagatgac agaagaggat tacaaggcac tgatcgatat catggatcgg gccagaaaaa 120  
 tggttgtttc gaaggaagac ggaagaatga aaaaagcggc tcaagaaacg taaagaaacg 180  
 cctgaaatga accggcccta tagtaagaat aggccggttg ttttgatttc tatgcagact 240  
 ctcccgggtg catttcgcga tccatatcag gatgccagat gagcgggtct tcccctttgt 300  
 cccgcgccat atcactacta acagttttta agttcatttt gttccaaaat tccgctgatt 360  
 tcattctcgg atttgtcgg atcggcattt tgaatgattt tgcaaattca accaaggctc 420  
 tcccgtatcc cctgttctgg tagcccgga gaacctcaag cttccacagc tccaaataat 480  
 cctggcgggt gtcaaaatag ggattcgatt tgccgttaac ttgatacaga ctcatctgtg 540  
 ctacaagttt atcgccaaaa taaatcccgt aaaaaggcga ggtgctgtca ttttcaataa 600  
 tattatcctg aagttcttca agcattgaaa gtcctgaat gccgtattct ttgaatttct 660  
 tgaattcttc cagcgtttta tagttgataa gcagacgttc tacctttgtc aaacaaatct 720  
 ccccctttgt tgtttctaca tatattgtaa acgctttatt taaaaaatcc aaatatttaa 780  
 actttaattt taagcacatg ggatctttga gaagtaattt cttcttaact ctgctatgat 840  
 aatacgtaaa tgcgtcgacc gaaacagcag ttataaggca tgaagctgtc cggtttttgc 900  
 aaaagtggct gtgactgtaa aaagaaatcg aaaaagaccg ttttgtgtga aaacggtctt 960



tttgtttcct ttttaaccaac tgccataact cgaggcctac cttagcttcca agaaagatat 1020  
cctaacagca caagagcgga aagatgtttt gttctacatc cagaacaacc tctgctaaaa 1080  
ttcctgaaaa attttgcaaa aagttgttga ctttatctac aaggtgtggt ataataatct 1140  
taacaacagc aggacgctct agaggaggag acatcatgaa aattggaatt atcggcggag 1200  
gctccgttgg tcttttatgc gcctattatt tgtcacttta tcacgacgtg actgttgtga 1260  
cgaggcggca agaacaggct gcggccattc agtctgaagg aatccggctt tataaaggcg 1320  
gggaggaatt cagggtgat tgcagtgcgg acacgagtat caattcggac tttgacctgc 1380  
ttgtcgtgac agtgaagcag catcagcttc aatctgtttt ttcgtcgctt gaacgaatcg 1440  
ggaagacgaa tatattattt ttgcaaaacg gcatggggca tatccacgac ctaaaagact 1500  
ggcacgttgg ccattccatt tatgttgga tctgtgagca cggagctgta agaaaatcgg 1560  
atacagctgt tgatcataca ggcttaggtg cgataaaatg gagcgcgttc gacgatgctg 1620  
aaccagaccg gctgaacatc ttgtttcagc ataaccattc ggattttccg atttattatg 1680  
agacggattg gtaccgtctg ctgacgggca agctgattgt aaatgcgtgt attaatacctt 1740  
taactgcgtt attgcaagtg aaaaatggag aactgctgac aacgccagct tatctggctt 1800  
ttatgaagct ggtatttcag gaggcattgcc gcattttaaa acttgaaaat gaagaaaagg 1860  
cttgggagcg ggttcaggcc gtttgtgggc aaacgaaaga gaatcgttca tcaatgctgg 1920  
ttgacgtcat tggaggccgg cagacggaag ctgacgccat tatcggatac ttattgaagg 1980  
aagcaagtct tcaaggctct gatgccgtcc acctagagtt tttatatggc agcatcaaag 2040  
cattggagcg aaatacaaac aaagtctttt gagctttttc ggtaacatgc tatactcatt 2100  
tcggatcact aactatttat tggagaaagg aagttctaga agatgcagct aactgaactt 2160  
tccatcaaaa atcagaatgt gtttgtacag cactatatag atggcaaaga agaaatgtct 2220  
tctttttttg attacagtat tcatcataag gacatgtggc gcgaaagact ggaagactta 2280  
tcttcccgggt ttttcgcaag agaggaattg gcggcgctact taacctctta ccataataaa 2340  
ttcggttcaa gtgcgatgca gtctgtatt gagaagctga aggaccgctc aagtgccgct 2400  
gtagtcggcg gacagcaggc aggactttta acaggaccgc ttacaccat acataaaatc 2460  
atttcaatca ttgttttagc aaagcaacaa gaaaaggaac tgcaagtgcc tgtcatacca 2520  
atcttctggg tggctggaga agaccacgat ttggatgaga ttaattttgt tcacacatct 2580  
gaagagaatg ggcctgtgaa aaaaaagctg cctcagtctt attggaagaa atcatcagca 2640  
gcgagtacat cgcttgatca ggaaaagtgt gccgcgtgga tagatgatgt ttttgccgct 2700

tttgaagaaa cagaccatac gaatacactt ctgcacaatg tgaaacgatg ttttaagggaa 2760  
tctgttacgt ttactgactt ctttgaactg ctgatcgcgg atttgttcca agaagagggc 2820  
ttagttttat taaattctgg ggatcctggt ttggcggatg agagaagatt ttcagcctga 2880  
tacagattaa atcagaacgc agaagcggtc tgataaaaca gaatttgcct ggcggcagta 2940  
gcgcggtggt cccacctgac cccatgccga actcagaagt gaaacgccgt agcgccgatg 3000  
gtagtgtggg gtctcccat gcgagagtag ggaactgcca ggcatacaat aaaacgaaag 3060  
gctcagtcga aagactgggc ctttcgtttt atctgttggt tgtcggtgaa cgctctcctg 3120  
agtaggacaa atccgccggg agcggatttg aacgttgcca agcaacggcc cggaggggtg 3180  
cgggcaggac gcccgccata aactgccagg catcaaatta agcagaaggc catcctgacg 3240  
gatggccttt ttgcgtttct acaaactctt ggtaccacaga aaaagcggca aaagcggctg 3300  
ttaaaaaagc gaaatcgaag aagctgtctg ccgctaagac ggaatatcaa aagcgttctg 3360  
ctgttggtgc atctttaaaa gtcacagccg atgaatccca gcaagatgtc ctaaaatact 3420  
tgaacaccca gaaagataaa ggaaatgcag accaaattca ttcttattat gtggtgaacg 3480  
ggattgctgt tcatgcctca aaagaggtta tggaaaaagt ggtgcagttt cccgaagtgg 3540  
aaaaggtgct tcctaatagag aaacggcagc tttttaagtc atcctcccca tttaatatga 3600  
aaaaagcaca gaaagctatt aaagcaactg acggtgtgga atggaatgta gaccaaatcg 3660  
atgccccaaa agcttgggca cttggatatg atggaactgg cacggttggt gcgtccattg 3720  
ataccggggt ggaatggaat catccggcat taaaagagaa atatcgcgga tataatccgg 3780  
aaaatcctaa tgagcctgaa aatgaaatga actggtatga tgccgtagca ggcgaggcaa 3840  
gcccttatga tgatttggct catggaaccc acgtgacagg cacgatgggt ggctctgaac 3900  
ctgatggaac aaatcaaact ggtgtagcac ctggcgcaaa atggattgct gttaaagcgt 3960  
tctctgaaga tggcggcact gatgctgaca ttttggaagc tggatgaatgg gttttagcac 4020  
caaaggacgc ggaaggaaat cccaccccg aaatggctcc tgatgttgct aataactcat 4080  
ggggaggggg ctctggactt gatgaatggt acagagacat ggtcaatgcc tggcgttcgg 4140  
ccgatatttt ccctgagttt tcagcgggga atacggatct ctttattccc ggcgggcctg 4200  
gttctatcgc aaatccggca aactatccag aatcgtttgc aactggagcg actgagaatt 4260  
ccaattcccc atggagagaa aagaaaatcg ctaatgttga ttactttgaa cttctgcata 4320  
ttcttgaatt taaaaaggct gaaagagtaa aagattgtgc tgaaatatta gagtataaac 4380  
aaaatcgtga aacaggcgaa agaaagtgtg atcgagtgtg gttttgtaaa tccaggcttt 4440

- 116 -

gtccaatgtg caactggagg agagcaatga aacatggcat tcagtcacaa aaggttgttg 4500  
ctgaagttat taaacaaaag ccaacagttc gttggttgtt tctcacatta acagttaaaa 4560  
atgtttatga tggcgaagaa ttaaataaga gtttgtcaga tatggctcaa ggatttcgcc 4620  
gaatgatgca atataaaaaa attaataaaa atcttgttgg ttttatgcgt gcaacggaag 4680  
tgacaataaa taataaagat aattcttata atcagcacat gcatgtattg gtatgtgtgg 4740  
aaccaactta ttttaagaat acagaaaact acgtgaatca aaaacaatgg attcaatttt 4800  
ggaaaaaggc aatgaaatta gactatgatc caaatgtaaa agttcaaatg attcgaccga 4860  
aaaataaata taaatcggat atacaatcgg caattgacga aactgcaaaa tctcctgtaa 4920  
aggatacggg ttttatgacc gatgatgaag aaaagaattt gaaacgtttg tctgatttgg 4980  
aggaagggtt acaccgtaaa aggttaatct cctatgggtg tttgttaaaa gaaatacata 5040  
aaaaattaaa ccttgatgac acagaagaag gcgatttgat tcatacagat gatgacgaaa 5100  
aagccgatga agatggattt tctattattg caatgtggaa ttgggaacgg aaaaattatt 5160  
ttattaaaga gtagttcaac aaacgggcca tattgttgta taagtgatga aatactgaat 5220  
ttaaaactta gtltatatgt ggtaaaatgt tttaatcaag tttaggagga attaattatg 5280  
aagtgtaatg aatgtaacag ggttcaatta aaagagggaa gcgtatcatt aaccctataa 5340  
actacgtctg ccttcattat tggagggtga aatgtgaata catcctattc acaatcgaat 5400  
ttacgacaca accaaatttt aatttggtt tgcattttat ctttttttag cgtattaaat 5460  
gaaatgggtt tgaacgtctc attacctgat attgcaaatg attttaataa accacctgcg 5520  
agtacaaact gggatgaacac agcctttatg ttaacctttt ccattggaac agctgtatat 5580  
ggaaagctat ctgatcaatt aggcacaaa aggttactcc tatttggaat tataataaat 5640  
tgtttcgggt cggttaattg gtttgttggc cattctttct tttccttact tattatggct 5700  
cgttttattc aaggggctgg tgcagctgca tttccagcac tcgtaatggg tgtagttgcg 5760  
cgctatattc caaaggaaaa taggggtaaa gcatttggtc ttattggatc gatagtagcc 5820  
atgggagaag gagtcggtcc agcgattggg ggaatgatag ccattatat tcattgggtcc 5880  
tatcttctac tcattcctat gataacaatt atcactgttc cgtttcttat gaaattatta 5940  
aagaaagaag taaggataaa aggtcatttt gatatcaaag gaattatact aatgtctgta 6000  
ggcattgtat tttttatgtt gtttacaaca tcatatagca tttcttttct tatcgtttagc 6060  
gtgctgtcat tcctgatatt tgtaaaacat atcaggaaaag taacagatcc ttttgttgat 6120  
cccggattag ggaaaaatat accttttatg attggagttc tttgtggggg aattatattt 6180

- 117 -

ggaacagtag caggggttgt ctctatggtt ccttatatga tgaaagatgt tcaccagcta 6240  
agtactgccg aaatcggaag tgtaattatt ttccctggaa caatgagtgt cattattttc 6300  
ggctacattg gtgggatact tgttgataga agaggtcctt tatacgtggt aaacatcgga 6360  
gttacatttc tttctgtag ctttttaact gcttcctttc ttttagaaac aacatcatgg 6420  
ttcatgacaa ttataatcgt atttgtttta ggtgggcttt cgttcaccaa aacagttata 6480  
tcaacaattg tttcaagtag ctgaaacag caggaagctg gtgctggaat gagtttgctt 6540  
aactttacca gctttttatc agagggaaca ggtattgcaa ttgtaggtgg tttattatcc 6600  
atacccttac ttgatcaaag gttgttacct atggaagttg atcagtcaac ttatctgtat 6660  
agtaatttgt tattactttt ttcaggaatc attgtcatta gttggctggg taccttgaat 6720  
gtatataaac attctcaaag ggattttctaa atcggttaagg gatcaacttt gggagagagt 6780  
tcaaaattga tccttttttt ataacagttc gaagcggccg caattcttga agacgaaagg 6840  
gcctcgtgat acgcctatct ttataggtta atgtcatgat aataatggtt tcttagacgt 6900  
caggtggcac ttttcgggga aatgtgcgcg gaacccttat ttgtttattt ttctaaatac 6960  
attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa 7020  
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat 7080  
tttgccttcc tgtttttgct caccagaaa cgctggtgaa agtaaaagat gctgaagatc 7140  
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga 7200  
gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg 7260  
cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc 7320  
agaatgactt gggtgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag 7380  
taagagaatt atgcagtgt gccataacca tgagtataa cactgcggcc aacttacttc 7440  
tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg 7500  
taactcgctt tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 7560  
acaccacgat gcctgcagca atggcaacaa cgttgcgcaa actattaact ggcgaactac 7620  
ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac 7680  
cacttctgag ctcggccctt ccggctgggt ggtttattgc tgataaatct ggagccggtg 7740  
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgatatc 7800  
tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 7860  
agataggtgc ctactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac 7920

tttagattga tttaaaactt catttttaat ttaaaaggat ctagggtgaag atcctttttg 7980  
ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg 8040  
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc 8100  
aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc 8160  
tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt 8220  
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc 8280  
taatcctggt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact 8340  
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggg tcgtgcacac 8400  
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 8460  
aaagcgccac gcttcccga gggagaaagg cggacaggta tccggtaagc ggcagggtcg 8520  
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcttg 8580  
tcgggtttcg ccacctctga cttgagcgtc gattttttgtg atgctcgtca ggggggcgga 8640  
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt 8700  
ttgctcacat gttcttttct gcgttatccc ctgattctgt ggataaccgt attaccgcct 8760  
ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg 8820  
aggaagcgga agagcgctg atgcggtatt ttctccttac gcactctgtgc ggtattttcac 8880  
accgcatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagtata 8940  
cactccgcta tcgctacgtg actgggtcat ggctgcgccc cgacaccgc caacaccgc 9000  
tgacgcgccc tgacgggctt gtctgtccc ggcacccgt tacagacaag ctgtgaccgt 9060  
ctccgggagc tgcattgtgc agaggttttc accgtcatca ccgaaacgcg cgaggcagct 9120  
gcggtaaagc tcatcagcgt ggtcgtgaag cgattcacag atgtctgcct gttcatccgc 9180  
gtccagctcg ttgagtttct ccagaagcgt taatgtctgg cttctgataa agcggggccat 9240  
gttaagggcg gtttttttct gtttggtcac ttgatgcctc cgtgtaaggg ggaattttctg 9300  
ttcatggggg taatgatacc gatgaaacga gagaggatgc tcacgatacg ggttactgat 9360  
gatgaacatg cccggttact ggaacgttgt gagggtaaac aactggcggg atggatgcgg 9420  
cgggaccaga gaaaaatcac tcagggtcaa tgccagcgt tcgttaatac agatgtaggt 9480  
gttccacagg gtagccagca gcactctgcg atgcagatcc ggaacataat ggtgcagggc 9540  
gctgacttcc gcgtttccag actttacgaa acacggaaac cgaagaccat tcatgttggt 9600  
gctcaggctc cagacgtttt gcagcagcag tcgcttcacg ttcgctcgcg tatcggtgat 9660

- 119 -

tcattctgct aaccagtaag gcaaccccgc cagcctagcc gggtcctcaa cgacaggagc 9720  
 acgatcatgc gcacccgtgg ccaggaccca acgctgcccg agatgcgccc cgtgcggctg 9780  
 ctggagatgg cggacgcgat ggatatgttc tgccaagggc tggtttgccg attcacagtt 9840  
 ctccgcaaga attgattggc tccaattctt ggagtgggtga atccgttagc gaggtgccgc 9900  
 cggcttccat tcaggctcag gtggcccggc tccatgcacc gcgacgcaac gcggggaggc 9960  
 agacaaggta tagggcggcg cctacaatcc atgccaaacc gttccatgtg ctcgccgagg 10020  
 cggcataaat cgccgtgacg atcagcggtc cagtgatcga agttaggctg gtaagagccg 10080  
 cgagcgatcc ttgaagctgt cctgatggc cgtcatctac ctgcctggac agcatggcct 10140  
 gcaacgcggg catcccgatg ccgccggaag cgagaagaat cataatgggg aaggccatcc 10200  
 agcctcgcgt cg 10212

&lt;210&gt; 82

&lt;211&gt; 10426

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
 pAN267 plasmid

&lt;400&gt; 82

aacaaaattc tccagtcctc acatcggttt gaaaggagga agcggaagaa tgaagtaaga 60  
 gggattttttg actccgaagt aagtcttcaa aaaatcaaat aaggagtgtc aagaatgttt 120  
 gcaaaacgat tcaaaacctc tttactgccg ttattcgctg gatTTTTatt gctgtttcat 180  
 ttggttctgg caggaccggc ggctgcgagt gctgaaacgg cgaacaaatc gaatgagctt 240  
 acagcacctg cgatcaaaag cggaaccatt cttcatgcat ggaattggc gttcaatacg 300  
 ttaaaacaca atatgaagga tattcatgat gcaggatata cagccattca gacatctccg 360  
 attaaccaag taaaggaagg gaatcaagga gataaaagca tgtcgaactg gtactggctg 420  
 tatcagccga catcgtatca aattggcaac cgttacttag gtactgaaca agaatttaaa 480  
 gaaatgtgtg cagccgctga agaatatggc ataaaggcca ttgttgacgc ggtcatcaat 540  
 cataccacca gtgattatgc cgcgatttcc aatgagggtta agagtattcc aaactggaca 600  
 catggaaaca cacaaattaa aaactggctc gatcgaaata gtacataatg gatttcctta 660  
 cgcgaaatac gggcagacat ggccctgccg gttattatta tttttgacac cagaccaact 720  
 ggtaatggta gcgaccggcg ctcaggatcg tctcggtacc aagagtttgt agaaacgcaa 780  
 aaaggccatc cgtcaggatg gccttctgct taatttgatg cctggcagtt tatggcgggc 840

gtcctgcccc ccaccctccg ggcggttget tcgcaacggt caaatccgct cccggcggat 900  
ttgtcctact caggagagcg ttcaccgaca aacaacagat aaaacgaaag gccaggtctt 960  
tcgactgagc ctttcgtttt atttgatgcc tggcagttcc ctactctcgc atggggagac 1020  
cccacactac catcggcgct acggcgtttc acttctgagt tcggcatggg gtcaggtggg 1080  
accaccgogc tactgccgcc aggc aaattc tgttttatca gaccgcttct gcgttctgat 1140  
ttaatctgta tcaggctgaa aatcttctct catccgcaa aacaggatcc atcacgaagc 1200  
gtcgatatga aaaaattaat ttgcgcaac ggagaccacc gcttccttct tcttgccctg 1260  
tttcacaaac ggcattcatt cacgaagctt tcttccact acttcgattt gatgttcgtt 1320  
ctcgcttgca ttgatagcgt tgaaacgagg acggtttact tggttttcga cgatccactc 1380  
ttttgcgaat gtaccgtttt ggatatcttt taatacttct ttcatagatt cttttacttt 1440  
agcgctccaca acgcgagggc ctgatacgaa atctcccccac tgtgctgtgt cagagattga 1500  
atatctcatt cctgcaagtc cttcttcgta cataaggtct acgattaatt tcagctcatg 1560  
aagacactcg aagtatgcaa gttcaggctg ataacctgct tcagttaagg tttcaaattcc 1620  
ggctttgaca agcgcgctta atccgcgcga aagaactgct tgctcaccga acaaactctgt 1680  
ttctgtttct tctttaaatg tcgtttctaa tacgcccgtc cttgcgcgcg cgattccttt 1740  
agcataagcg agggctttgt ctcttgcttc tccagtcaca tcttgataga ttgcgaacaa 1800  
tgcagggtacg ccagctcctt gctcatatgt tcttcttacc aagtgtcccg ggcctttagg 1860  
ggccactaag aatacatcta catccgcgcg aggaacaatt tgatggaaat gcacgttaaa 1920  
tccatgagcg aatactaattg attttcctgc tgtcaattca tctttgattt cagcttcgta 1980  
tactttttgc tgctgctcat ccggaagcag aaccatgatg atttcggctt gggctgccgc 2040  
ttcttttact gaaaatactt tatgtccgtc ttcttgggct tgagtgaag attttccttg 2100  
tetaacaccg acgatcacgt ctactccgct ttctttaagg ttcagggcat gtgcgtggcc 2160  
ttgcgaaccg tacccgataa ccgctactgt ttttcagcc aatacgttct ctttgatata 2220  
accgttataa tatactttta ccatttcaat ctctcccttg ttatgtttta tacaatagat 2280  
attgttttat tggatgacgc cttttgctgg ttcccctcgc aaaagccggt gtacctgttc 2340  
tcgcgatttc tttaatgcca taaggtttta ataactcaat aagcgcttca attttgtag 2400  
attcacctgt cacctgaaca acgatgctgt ctctgctgac atcaacgaca gaggctctaa 2460  
acggttctat gattccatta atctctgttc ttgttgaagg tgcggagaca accttgatta 2520  
aggccagctc cctctggaca atcgattgat ttgtgatgtc tgtgactttc agcacatcaa 2580

tctgtttggt gagctgtttc gtttaactggt caacatcatt ttcaccttca acatgaacga 2640  
cgaaggatgat tctggaaacg cgggctgttt ctgtgtgtcc aactgtaatg ctttcaatgt 2700  
tgtaatgcct ttttgtgaat agaccggtga tccggtttaa caccocggag cggttcacca 2760  
cagtcaatgt gataattctt ttcaagggtt caccocccacc atttcatgca gccctttccc 2820  
cggagccacc atcgggaata ctttttcttc gctggcaacc cgcacgtcaa tgacaacagg 2880  
ttctcttgat gttaatgcct cttccagctt ttcctttgct tccgcttctg atgaaattct 2940  
gatgccttta atgccgtatg cttcggacaa tttgacgaag tcaggctgag aagcgaattt 3000  
agattctgaa taacgttctt catagaaaat ttctgtccac tgtctgacca ttccgagaca 3060  
agcgttattt aaaatcacta ccttgaccgg aagatttaat tcgcgaataa catcgagttc 3120  
ttgaagcgtc atttggaatc cgcgctctcc gacaaccgag acaacagtag catctttttc 3180  
ggccagctgt gcgccgatcg ccgcgggaag accgaatccc atcgttccaa gtccgcctga 3240  
cgtgacccat ttatctgctt tttggaacgg ataaaattgc gctgaccaca tttgatgctg 3300  
gcctacatcc gttgcgacaa tggcctctcc ttttgtaaatt tgatgaatat attcaatcaa 3360  
tttctgaggt ttaaaacctt cttcttcatt atctacatac cagagcggat actcttcttt 3420  
ccattctgog agctgttttt tccattcgct tgaatcgctt tgtttgccgt cttgtttgat 3480  
cagctcctgc aggacaattt tgctgtctcc gactacagga atctgtgttt tcatgatttt 3540  
tccgatttca gctggatcaa tatcgatgtg ggctatcttt gcgtttctgg caaagtgttt 3600  
caggtttcct gtgacacggt catcaaaacg ggcccgata ctgattaata gatcacattc 3660  
atgaagggcc atattggctg tataagtacc gtgcattccc gccatcccta ggaaaagcgg 3720  
atggtcagcc gggaagcctc cgagccccaa aagggtgtgt gccacaggga tttgctgctg 3780  
ttcagcataa ttttttaatt cttctgacgc ttttccgtgc agtacgcccg caccgcccag 3840  
gatcacgggt tttttcgac tgctcacggc ttccacaagc ttgcggatct gcaaataatt 3900  
cggctctggt gtcggctggt atcccgagg atctatctca tgatcgtagc tgaattctcc 3960  
ttcaattggt gctacatctt tcggaatatc aatcaatata ggtccgggtc ttccagttgt 4020  
tgcaatatgg aacgcttctt taatgatgog cggcagatct tccggctggc gaacctggta 4080  
gctgtgtttt gttactggca tcgtaatccc taaaatgtct gcttcctgaa atgcatcgct 4140  
cccgattaca gaggttgcta cctgcctgt aaagacgact aacggcaatg aatcaatcat 4200  
ggcatcagca aggctgttaa caaggtttgt cgctcccggc ctgacgtggc aatgacgaca 4260  
ccggtttccg gagacccttg cgtatccctc cgctgcatga attgctcctt gttcgtgacg 4320



ggaaggatat gtaccaacct gaatgtatag cttatcgtaa atcggaagca cagccccgcc 4380  
cggataaccg aagatcattt ctactttctc tttttttaat gattcaatca gcattaatcg 4440  
tccgctcacc gtctgtgtac attcggcaga tgctgaatcc acctgtacat tagtccccat 4500  
tttatctcct cctctagagc gtcctgctgt tggtaagatt attataccac accttgtaga 4560  
taaagtcaac aactttttgc aaaatttttc aggaatttta gcagaggttg ttctggatgt 4620  
agaacaaaac atctttccgc tcttgtgctg ttaggatatc tttcttgaa gctaggtagg 4680  
cctcgagtta tggcagttgg ttaaaaggaa acaaaaagac cgttttcaca caaaacggtc 4740  
tttttcgatt tctttttaca gtcacagcca cttttgcaaa aaccggacag cttcatgcct 4800  
tataactgct gtttcggtcg acgaagcggc cgccgtttta acgaattcct gcagctggcg 4860  
aatggcgatt ttcgttcgtg aatacatgtt ataataacta taactaataa cgtaacgtga 4920  
ctggcaagag atatttttaa aacaatgaat aggtttacac ttactttagt tttatggaaa 4980  
tgaaagatca tatcatatat aatctagaat aaaattaact aaaataatta ttatctagat 5040  
aaaaaattta gaagccaatg aaatctataa ataaactaaa ttaagtttat ttaattaaca 5100  
actatggata taaaataggt actaatcaaa atagtgagga ggatatattt gaatacatat 5160  
gaacaaatta ataaagtga aaaaatactt cggaacatt taaaaataa ccttattggt 5220  
acttacatgt ttggatcagg agttgagagt ggactaaaac caaatagtga tcttgacttt 5280  
ttagtcgtcg tatctgaacc attgacagat caaagtaaag aaatacttat acaaaaaatt 5340  
agacctattt caaaaaaat aggagataaa agcaacttac gatatatatga attaacaatt 5400  
attattcagc aagaaatggt accgtggaat catcctccca aacaagaatt tatttatgga 5460  
gaatggttac aagagcttta tgaacaagga tacattcctc agaaggaatt aaattcagat 5520  
ttaaccataa tgctttacca agcaaacga aaaaataaaa gaatatacgg aaattatgac 5580  
ttagaggaat tactacctga tattccattt tctgatgtga gaagagccat tatggattcg 5640  
tcagaggaat taatagataa ttatcaggat gatgaaacca actctatatt aactttatgc 5700  
cgtatgattt taactatgga cacgggtaaa atcataccaa aagatatatgc gggaaatgca 5760  
gtggctgaat cttctccatt agaacatagg gagagaattt tgtagcagt tcgtagttat 5820  
cttgagaga atattgaatg gactaatgaa aatgtaaatt taactataaa ctatttaaatt 5880  
aacagattaa aaaaattata aaaaaattga aaaaatggtg gaaacacttt tttcaatttt 5940  
tttgttttat tatttaatat ttgggaaata ttcattctaa ttggtaatca gattttagaa 6000  
aacaataaac ccttgcatag ggggatcgat atccgttttag gctgggcggt gatagcttct 6060

cgttcaggca gtacgcctct tttcttttcc agacctgagg gaggcggaaa tgggtgtgagg 6120  
ttcccgggga aaagccaaat aggcgatcgc gggagtgcctt tatttgaaga tcaggctatc 6180  
actgcggtca atagatttca caatgtgatg gctggacagc ctgaggaact ctogaacccg 6240  
aatggaaaca accagatatt tatgaatcag cgcggtcac atggcgttgt gctggcaa at 6300  
gcaggttcat cctctgtctc tatcaatacg gcaacaaaat tgcctgatgg caggatgac 6360  
aataaagctg gagcgggttc atttcaagtg aacgatggta aactgacagg cacgatcaat 6420  
gccaggctctg tagctgtgct ttatcctgat gatattgcaa aagcgctca tgttttcctt 6480  
gagaattaca aaacaggtgt aacacattct ttcaatgatc aactgacgat taccttgcgt 6540  
gcagatgcga atacaacaaa agccgtttat caaatcaata atggaccaga cgacaggcgt 6600  
ttaaggatgg agatcaattc acaatcgga aaggagatcc aatttggcaa aacatacacc 6660  
atcatgttaa aaggaacgaa cagtgatggg gtaacgagga ccgagaaata cagttttgtt 6720  
aaaagagatc cagcgtcggc caaaaccatc ggctatcaaa atccgaatca ttggagccag 6780  
gtaaatgctt atatctataa acatgatggg agccgagtaa ttgaattgac cggatcttgg 6840  
cctggaaaac caatgactaa aaatgcagac ggaatttaca cgctgacgct gcctgcggac 6900  
acggatacaa ccaacgcaaa agtgattttt aataatggca gcgccaagt gcccggtcag 6960  
aatcagcctg gctttgatta cgtgctaaat ggtttatata atgactcggg ctttaagcgtt 7020  
tctcttcccc attgagggca aggctagacg ggacttaccg aaagaaacca tcaatgatgg 7080  
tttctttttt gttcataaat cagacaaaac ttttctcttg caaaagtttg tgaagtgttg 7140  
cacaatataa atgtgaaata cttcaciaac aaaaagacat caaagagaaa cataccctgc 7200  
aaggatgctg atattgtctg catttgcgcc ggagcaaacc aaaaacctgg tgagacacgc 7260  
cttgaattag tagaaaagaa cttgaagatt ttcaaaggca tcgttagtga agtcatggcg 7320  
agcggatttg acggcatttt cttagtcggg cggcacctcg ctaacggatt caccactcca 7380  
agaattggag ccaatcaatt cttgcggaga actgtgaatg cgcaaacc aa cccttggcag 7440  
aacatatcca tcgcgtccgc catctccagc agccgcacgc ggcgatctc gggcagcgtt 7500  
gggtcctggc cacgggtgcg catgatcgtg ctctgtcgt tgaggacccg gctaggctgg 7560  
cggggttgcc ttactggtta gcagaatgaa tcaccgatac gcgagcgaac gtgaagcgac 7620  
tgctgctgca aaacgtctgc gacctgagca acaacatgaa tgggtcttcgg tttccgtgtt 7680  
tcgtaaagtc tggaacgcg gaagtcagcg ccctgcacca ttatgttccg gatctgcac 7740  
gcaggatgct gctggctacc ctgtggaaca cctacatctg tattaacgaa gcgctggcat 7800

tgaccctgag tgatttttct ctggtcccgc cgcattccata ccgccagttg tttaccctca 7860  
caacgttcca gtaaccgggc atgttcatca tcagtaaccc gtatcgtgag catcctctct 7920  
cgtttcatcg gtatcattac ccccatgaac agaaatcccc cttacacgga ggcattcagt 7980  
accaaacagg aaaaaaccgc ccttaacatg gcccgcttta tcagaagcca gacattaacg 8040  
cttctggaga aactcaacga gctggacgcg gatgaacagg cagacatctg tgaatcgctt 8100  
cacgaccacg ctgatgagct ttaccgcagc tgcctcgcgc gtttcggtga tgacggtgaa 8160  
aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg 8220  
agcagacaag cccgtcaggg cgcgtcagcg ggtgttgccg ggtgtcgggg cgcagccatg 8280  
accagtcac gtagcgatag cggagtgtat actggcttaa ctatgcggca tcagagcaga 8340  
ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta aggagaaaat 8400  
accgcatcag gcgctcttcc gcttctcgc tcaactgactc gctgcgctcg gtcgttcggc 8460  
tgccggcagc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 8520  
ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 8580  
ccgcgttgct ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac 8640  
gctcaagtca gaggtggcga aaccgcagc gactataaag ataccaggcg tttccccctg 8700  
gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 8760  
ttctcccttc ggggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcaagtccg 8820  
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 8880  
gcgccttctc cggttaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac 8940  
tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 9000  
tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttgggt atctgcgctc 9060  
tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 9120  
ccgctggtag cgggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 9180  
ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 9240  
gttaagggat tttggtcacg agattatcaa aaaggatctt cacctagatc cttttaaatt 9300  
aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggctc gacagttacc 9360  
aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg 9420  
cctgactccc cgtcgtgtag ataactacga tacggggagg cttaccatct ggccccagtg 9480  
ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc 9540

- 125 -

cagccggaag ggccgagcgc agaagtgggc ctgcaacttt atccgcctcc atccagtcta 9600  
 ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg 9660  
 ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt tggatatggt tcattcagct 9720  
 ccggttccca acgatcaagg cgagttacat gatcccccac gttgtgcaaa aaagcgggta 9780  
 gtccttcggg tcctccgacg gttgtcagaa gtaagttggc cgcagtgtta tcaatcatgg 9840  
 ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga 9900  
 ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 9960  
 gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 10020  
 ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 10080  
 cgatgtacc cactcgtgca cccaactgat ctcagcatc ttttactttc accagcgttt 10140  
 ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa gggaataagg gcgacacgga 10200  
 aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat cagggttatt 10260  
 gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 10320  
 gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa 10380  
 cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaatt 10426

&lt;210&gt; 83

&lt;211&gt; 4191

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
 pAN263 plasmid

&lt;400&gt; 83

ccaatacgca aaccgcctct cccgcgcgt tggccgattc attaatgcag cccggatccg 60  
 aaagaagagg atatcccgtt ttacagaag gcattggatg atccaaagggt gtccatcaga 120  
 agacaggctg ttgtgtactt aggaatgatt gaaacacctg atgttcttcc tctattgtat 180  
 aaagcacttg aggacaaagc tgtatcagtc agaagaacgg ccggagactg cctgtctgat 240  
 atcggcgatc ctcaagccat tcctgctatg atcaagtcac taagcgactc cagcaagctt 300  
 gttcgctggc gtgccgccat gttcctgtac gaagtcggcg atgaaagtgc aattgaagct 360  
 ttgcgcgctg ccgaagatga cccgaattt gaggtcagcc ttcaagtcaa aatggcgctt 420  
 gaacgtattg agcatggaga agaagcaaaa gggttctgtt ggaaacaaat gacggaaagc 480

- 126 -

agaaaaaaag gcgaataaag ataaaaaaagg tgcagatcat gcaccttttt tatgtgaatt 540  
ggtcgaccga aacagcagtt ataaggcatg aagctgtccg gtttttgcaa aagtggctgt 600  
gactgtaaaa agaaatcgaa aaagaccgtt ttgtgtgaaa acggtctttt tgtttccttt 660  
taaccaactg ccataactcg aggctacct agcttccaag aaagatatcc taacagcaca 720  
agagcggaaa gatgttttgt tctacatcca gaacaacctc tgctaaaatt cctgaaaaat 780  
tttgcaaaaa gttgttgact ttatctacaa ggtgtggtat aataatctta acaacagcag 840  
gacgctctag aggaggagac accatggcag aattacgcag taatatgac acacaaggaa 900  
tcgatagagc tccgcaccgc agtttgcttc gtgcagcagg ggtaaaagaa gaggatttcg 960  
gcaagccgtt tattgcggtg tgtaattcat acattgatat cgttcccgtt catgttcact 1020  
tgcaggagtt tgggaaaatc gtaaaagaag caatcagaga agcagggggc gttccgtttg 1080  
aatttaatac cattggggta gatgatggca tcgcaatggg gcatatcggg atgagatatt 1140  
cgctgccaaag ccgtgaaatt atcgcagact ctgtggaaac ggttgatcc gcacactggg 1200  
ttgacggaat ggtctgtatt ccgaactgcg acaaaatcac accgggaatg cttatggcgg 1260  
caatgcgcat caacattccg acgatttttg tcagcggcgg accgatggcg gcaggaagaa 1320  
caagttacgg gcgaaaaatc tccctttcct cagtattcga aggggtaggc gcctaccaag 1380  
cagggaaaat caacgaaaac gagcttcaag aactagagca gttcggatgc ccaacgtgcg 1440  
ggctcttgctc aggcattgtt acggcgaaact caatgaactg tctgtcagaa gcacttggtc 1500  
ttgctttgcc gggtaatgga accattctgg caacatctcc ggaacgcaaa gagtttgtga 1560  
gaaaatcggc tgcgcaatta atggaaacga ttcgcaaaga tatcaaaccg cgtgatattg 1620  
ttacagtaaa agcgattgat aacgcgtttg cactcgatat ggcgctcgga ggttctacaa 1680  
atacgtttct tcataccctt gcccttgcaa acgaagccgg cgttgaatac tctttagaac 1740  
gcattaacga agtcgctgag cgcgtgccgc acttggttaa gctggcgctt gcacgggatg 1800  
tgtttattga agatcttcac gaagcggggc gcgtttcagc ggctctgaat gagctttcga 1860  
agaaagaagg agcgcttcat ttagatgcgc tgactgttac aggaaaaact cttggagaaa 1920  
ccattgccgg acatgaagta aaggattatg acgtcattca cccgctggat caaccattca 1980  
ctgaaaaggg aggccttgct gttttattcg gtaatctagc tccggacggc gctatcatta 2040  
aaacaggcgg cgtacagaat gggattacaa gacacgaagg gccggctgtc gtattcgatt 2100  
ctcaggacga ggcgcttgac ggcattatca accgaaaagt aaaagaaggc gacgttgtca 2160  
tcacagata cgaagggcc aaggcggac ctggcatgcc ggaaatgctg gcgccaacat 2220

cccaaactcgt tgggaatggga ctcggggccaa aagtggcatt gattacggac ggacgttttt 2280  
ccggagcctc ccgtggcctc tcaatcggcc acgtatcacc tgaggccgct gagggcgggc 2340  
cgcttgccctt tgttgaaaac ggagaccata ttatcggtga tattgaaaaa cgcattcttg 2400  
atgtacaagt gccagaagaa gagtgggaaa aacgaaaagc gaactggaaa ggttttgaac 2460  
cgaaagtga aaccggctac ctggcacgtt attctaaact tgtgacaagt gccaacaccg 2520  
gcggtattat gaaaatctag acccctggcg taatagcgaa gagggccgca ccgatcgccc 2580  
ttccaacag ttgcgcagcc tgaatggcg atgagcttgc gccgtcccgt caagtcagcg 2640  
taatgctctg ccagtgttac aaccaattaa ccaattctga ttagaaaaac tcatcgagca 2700  
tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatttt tgaaaaagcc 2760  
gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca agatcctggt 2820  
atcggctctgc gattccgact cgtccaacat caatacaacc tattaatttc cctcgtcaa 2880  
aaataagggt atcaagtgag aaatcaccat gagtgacgac tgaatccggt gagaatggca 2940  
aaaggttatg catttctttc cagacttggt caacaggcca gccattacgc tegtcatcaa 3000  
aatcactcgc atcaacaaaa ccgttattca ttcgtgattg cgcctgagcg agacgaaata 3060  
cgcgatcgct gttaaaagga caattacaaa caggaatcga atgcaaccgg cgcaggaaca 3120  
ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttctaata acctggaatg 3180  
ctgttttccc agggatcgca gtggtgagta accatgcac atcaggagta cggataaaat 3240  
gcttgatggt cggaagaggc ataaattccg tcagccagtt tagtctgacc atctcatctg 3300  
taacatcatt ggcaacgcta cctttgccat gtttcagaaa caactctggc gcacggggct 3360  
tcccatataa tcaatagatt gtcgcacctg attgcccgc attatcgca gccatttat 3420  
acccatataa atcagcatcc atgttggaat ttaatcgcg cctcgacgag caagacgttt 3480  
cccgttgaat atggctcata acacccttg tattactgtt tatgtaagca gacagtttta 3540  
ttgttcatga tgatatattt ttatcttggt caatgtaaca tcagagattt tgagacactc 3600  
gacaagatga tcttcttgag atcgttttgg tctgcgcgta atctcttgct ctgaaaacga 3660  
aaaaaccgcc ttgcaggcg gtttttcgaa ggttctctga gctaccaact ctttgaaccg 3720  
aggtaactgg cttggaggag cgcagtcacc aaaacttgct ctttcagttt agccttaacc 3780  
ggcgcacgac ttcaagacta actcctctaa atcaattacc agtggctgct gccagtgggt 3840  
cttttgcacg tctttccggg ttggactcaa gacgatagtt accggataag gcgcagcggt 3900  
cggactgaac ggggggttcg tgcatacagt ccagcttga gcaactgcc taccgggaac 3960

- 128 -

tgagtgtcag gcgtggaatg agacaaacgc ggccataaca gcggaatgac accggtaaac 4020  
 cgaaaggcag gaacaggaga gcgcacgagg gagccgccag gggaaacgcc tggtatcttt 4080  
 atagtctgt cgggtttcgc caccactgat ttgagcgtca gatttcgtga tgcttgtcag 4140  
 gggggcggag cctatggaaa aacggctttg ccgcgccct ctcacttccc t 4191

&lt;210&gt; 84

&lt;211&gt; 702

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(699)

&lt;400&gt; 84

ttg tta ctg gtt atc gat gtg ggg aac acc aat act gta ctt ggt gta 48  
 Met Leu Leu Val Ile Asp Val Gly Asn Thr Asn Thr Val Leu Gly Val  
 1 5 10 15

tat cat gat gga aaa tta gaa tat cac tgg cgt ata gaa aca agc agg 96  
 Tyr His Asp Gly Lys Leu Glu Tyr His Trp Arg Ile Glu Thr Ser Arg  
 20 25 30

cat aaa aca gaa gat gag ttt ggg atg att ttg cgc tcc tta ttt gat 144  
 His Lys Thr Glu Asp Glu Phe Gly Met Ile Leu Arg Ser Leu Phe Asp  
 35 40 45

cac tcc ggg ctt atg ttt gaa cag ata gat ggc att att att tcg tca 192  
 His Ser Gly Leu Met Phe Glu Gln Ile Asp Gly Ile Ile Ile Ser Ser  
 50 55 60

gta gtg ccg cca atc atg ttt gcg tta gaa aga atg tgc aca aaa tac 240  
 Val Val Pro Pro Ile Met Phe Ala Leu Glu Arg Met Cys Thr Lys Tyr  
 65 70 75 80

ttt cat atc gag cct caa att gtt ggt cca ggt atg aaa acc ggt tta 288  
 Phe His Ile Glu Pro Gln Ile Val Gly Pro Gly Met Lys Thr Gly Leu  
 85 90 95

aat ata aaa tat gac aat ccg aaa gaa gta ggg gca gac aga atc gta 336  
 Asn Ile Lys Tyr Asp Asn Pro Lys Glu Val Gly Ala Asp Arg Ile Val  
 100 105 110

aat gct gtc gct gcg ata cac ttg tac ggc aat cca tta att gtt gtc 384  
 Asn Ala Val Ala Ala Ile His Leu Tyr Gly Asn Pro Leu Ile Val Val  
 115 120 125

gat ttc gga acc gcc aca acg tac tgc tat att gat gaa aac aaa caa 432  
 Asp Phe Gly Thr Ala Thr Thr Tyr Cys Tyr Ile Asp Glu Asn Lys Gln  
 130 135 140

tac atg ggc ggg gcg att gcc cct ggg att aca att tcg aca gag gcg 480  
 Tyr Met Gly Gly Ala Ile Ala Pro Gly Ile Thr Ile Ser Thr Glu Ala  
 145 150 155 160

- 129 -

ctt tac tcg cgt gca gca aag ctt cct cgt atc gaa atc acc cgg ccc 528  
 Leu Tyr Ser Arg Ala Ala Lys Leu Pro Arg Ile Glu Ile Thr Arg Pro  
 165 170 175

gac aat att atc gga aaa aac act gtt agc gcg atg caa tct gga att 576  
 Asp Asn Ile Ile Gly Lys Asn Thr Val Ser Ala Met Gln Ser Gly Ile  
 180 185 190

tta ttt ggc tat gtc ggc caa gtg gaa gga atc gtt aag cga atg aaa 624  
 Leu Phe Gly Tyr Val Gly Gln Val Glu Gly Ile Val Lys Arg Met Lys  
 195 200 205

tgg cag gca aaa cag gac cca agg tca ttg cga cag gag gcc tgg cgc 672  
 Trp Gln Ala Lys Gln Asp Pro Arg Ser Leu Arg Gln Glu Ala Trp Arg  
 210 215 220

cgc tca ttg cga acg aat cag att gta tag 702  
 Arg Ser Leu Arg Thr Asn Gln Ile Val  
 225 230

&lt;210&gt; 85

&lt;211&gt; 233

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 85

Met Leu Leu Val Ile Asp Val Gly Asn Thr Asn Thr Val Leu Gly Val  
 1 5 10 15

Tyr His Asp Gly Lys Leu Glu Tyr His Trp Arg Ile Glu Thr Ser Arg  
 20 25 30

His Lys Thr Glu Asp Glu Phe Gly Met Ile Leu Arg Ser Leu Phe Asp  
 35 40 45

His Ser Gly Leu Met Phe Glu Gln Ile Asp Gly Ile Ile Ile Ser Ser  
 50 55 60

Val Val Pro Pro Ile Met Phe Ala Leu Glu Arg Met Cys Thr Lys Tyr  
 65 70 75 80

Phe His Ile Glu Pro Gln Ile Val Gly Pro Gly Met Lys Thr Gly Leu  
 85 90 95

Asn Ile Lys Tyr Asp Asn Pro Lys Glu Val Gly Ala Asp Arg Ile Val  
 100 105 110

Asn Ala Val Ala Ala Ile His Leu Tyr Gly Asn Pro Leu Ile Val Val  
 115 120 125

Asp Phe Gly Thr Ala Thr Thr Tyr Cys Tyr Ile Asp Glu Asn Lys Gln  
 130 135 140

Tyr Met Gly Gly Ala Ile Ala Pro Gly Ile Thr Ile Ser Thr Glu Ala  
 145 150 155 160



- 130 -

Leu Tyr Ser Arg Ala Ala Lys Leu Pro Arg Ile Glu Ile Thr Arg Pro  
 165 170 175  
 Asp Asn Ile Ile Gly Lys Asn Thr Val Ser Ala Met Gln Ser Gly Ile  
 180 185 190  
 Leu Phe Gly Tyr Val Gly Gln Val Glu Gly Ile Val Lys Arg Met Lys  
 195 200 205  
 Trp Gln Ala Lys Gln Asp Pro Arg Ser Leu Arg Gln Glu Ala Trp Arg  
 210 215 220  
 Arg Ser Leu Arg Thr Asn Gln Ile Val  
 225 230

&lt;210&gt; 86

&lt;211&gt; 1623

&lt;212&gt; DNA

&lt;213&gt; Bacillus subtilis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1620)

&lt;400&gt; 86

atg tat ttg gca ttc cag gtg caa aaa ttg atg cgg tat ttg acg ctt 48  
 Met Tyr Leu Ala Phe Gln Val Gln Lys Leu Met Arg Tyr Leu Thr Leu  
 1 5 10 15  
 tac aag ata aag gac ctg aaa tta tgc ttg ccc ggc acg aac aaa acg 96  
 Tyr Lys Ile Lys Asp Leu Lys Leu Ser Leu Pro Gly Thr Asn Lys Thr  
 20 25 30  
 cag caa ttc atg gcc caa gca gtc ggc cgt tta act gga aaa ccg gga 144  
 Gln Gln Phe Met Ala Gln Ala Val Gly Arg Leu Thr Gly Lys Pro Gly  
 35 40 45  
 gtc gtg tta gtc aca tca gga ccg ggt gcc tct aac ttg gca aca ggc 192  
 Val Val Leu Val Thr Ser Gly Pro Gly Ala Ser Asn Leu Ala Thr Gly  
 50 55 60  
 ctg ctg aca gcg aac act gaa gga gac cct gtc gtt gcg ctt gct gga 240  
 Leu Leu Thr Ala Asn Thr Glu Gly Asp Pro Val Val Ala Leu Ala Gly  
 65 70 75 80  
 aac gtg atc cgt gca tat cgt tta aaa cgg aca cat caa tct ttg gat 288  
 Asn Val Ile Arg Ala Tyr Arg Leu Lys Arg Thr His Gln Ser Leu Asp  
 85 90 95  
 aat gcg gcg cta ttc cag ccg att aca aaa tac agt gta gaa gtt caa 336  
 Asn Ala Ala Leu Phe Gln Pro Ile Thr Lys Tyr Ser Val Glu Val Gln  
 100 105 110  
 gat gta aaa aat ata ccg gaa gct gtt aca aat gca ttt agg ata gcg 384  
 Asp Val Lys Asn Ile Pro Glu Ala Val Thr Asn Ala Phe Arg Ile Ala  
 115 120 125  
 tca gca ggg cag gct ggg gcc gct ttt gtg agc ttt ccg caa gat gtt 432

- 131 -

Ser	Ala	Gly	Gln	Ala	Gly	Ala	Ala	Phe	Val	Ser	Phe	Pro	Gln	Asp	Val	
130						135					140					
gtg	aat	gaa	gtc	aca	aat	acg	aaa	aac	gtg	cgt	gct	gtt	gca	gcg	cca	480
Val	Asn	Glu	Val	Thr	Asn	Thr	Lys	Asn	Val	Arg	Ala	Val	Ala	Ala	Pro	
145					150					155					160	
aaa	ctc	ggg	cct	gca	gca	gat	gat	gca	atc	agt	gcg	gcc	ata	gca	aaa	528
Lys	Leu	Gly	Pro	Ala	Ala	Asp	Asp	Ala	Ile	Ser	Ala	Ala	Ile	Ala	Lys	
				165					170					175		
atc	caa	aca	gca	aaa	ctt	cct	gtc	gtt	ttg	gtc	ggc	atg	aaa	ggc	gga	576
Ile	Gln	Thr	Ala	Lys	Leu	Pro	Val	Val	Leu	Val	Gly	Met	Lys	Gly	Gly	
			180					185					190			
aga	ccg	gaa	gca	att	aaa	gcg	gtt	cgc	aag	ctt	ttg	aaa	aag	gtt	cag	624
Arg	Pro	Glu	Ala	Ile	Lys	Ala	Val	Arg	Lys	Leu	Leu	Lys	Lys	Val	Gln	
		195					200					205				
ctt	cca	ttt	gtt	gaa	aca	tat	caa	gct	gcc	ggg	acc	ctt	tct	aga	gat	672
Leu	Pro	Phe	Val	Glu	Thr	Tyr	Gln	Ala	Ala	Gly	Thr	Leu	Ser	Arg	Asp	
	210					215					220					
tta	gag	gat	caa	tat	ttt	ggc	cgt	atc	ggg	ttg	ttc	cgc	aac	cag	cct	720
Leu	Glu	Asp	Gln	Tyr	Phe	Gly	Arg	Ile	Gly	Leu	Phe	Arg	Asn	Gln	Pro	
225					230					235					240	
ggc	gat	tta	ctg	cta	gag	cag	gca	gat	gtt	gtt	ctg	acg	atc	ggc	tat	768
Gly	Asp	Leu	Leu	Leu	Glu	Gln	Ala	Asp	Val	Val	Leu	Thr	Ile	Gly	Tyr	
				245					250					255		
gac	ccg	att	gaa	tat	gat	ccg	aaa	ttc	tgg	aat	atc	aat	gga	gac	cgg	816
Asp	Pro	Ile	Glu	Tyr	Asp	Pro	Lys	Phe	Trp	Asn	Ile	Asn	Gly	Asp	Arg	
			260					265					270			
aca	att	atc	cat	tta	gac	gag	att	atc	gct	gac	att	gat	cat	gct	tac	864
Thr	Ile	Ile	His	Leu	Asp	Glu	Ile	Ile	Ala	Asp	Ile	Asp	His	Ala	Tyr	
		275					280					285				
cag	cct	gat	ctt	gaa	ttg	atc	ggg	gac	att	ccg	tcc	acg	atc	aat	cat	912
Gln	Pro	Asp	Leu	Glu	Leu	Ile	Gly	Asp	Ile	Pro	Ser	Thr	Ile	Asn	His	
	290					295					300					
atc	gaa	cac	gat	gct	gtg	aaa	gtg	gaa	ttt	gca	gag	cgt	gag	cag	aaa	960
Ile	Glu	His	Asp	Ala	Val	Lys	Val	Glu	Phe	Ala	Glu	Arg	Glu	Gln	Lys	
305					310					315					320	
atc	ctt	tct	gat	tta	aaa	caa	tat	atg	cat	gaa	ggg	gag	cag	gtg	cct	1008
Ile	Leu	Ser	Asp	Leu	Lys	Gln	Tyr	Met	His	Glu	Gly	Glu	Gln	Val	Pro	
				325					330					335		
gca	gat	tgg	aaa	tca	gac	aga	gcg	cac	cct	ctt	gaa	atc	gtt	aaa	gag	1056
Ala	Asp	Trp	Lys	Ser	Asp	Arg	Ala	His	Pro	Leu	Glu	Ile	Val	Lys	Glu	
			340					345					350			
ttg	cgt	aat	gca	gtc	gat	gat	cat	gtt	aca	gta	act	tgc	gat	atc	ggg	1104
Leu	Arg	Asn	Ala	Val	Asp	Asp	His	Val	Thr	Val	Thr	Cys	Asp	Ile	Gly	
		355					360					365				

- 132 -

tcg cac tcc att tgg atg tca cgt tat ttc cgc agc tac gag ccg tta	1152
Ser His Ser Ile Trp Met Ser Arg Tyr Phe Arg Ser Tyr Glu Pro Leu	
370 375 380	
aca tta atg atc agt aac ggt atg caa aca ctc ggc gtt gcg ctt cct	1200
Thr Leu Met Ile Ser Asn Gly Met Gln Thr Leu Gly Val Ala Leu Pro	
385 390 395 400	
tgg gca atc ggc gct tca ttg gtg aaa ccg gga gaa aaa gtg gtt tct	1248
Trp Ala Ile Gly Ala Ser Leu Val Lys Pro Gly Glu Lys Val Val Ser	
405 410 415	
gtc tct ggt gac ggc ggt ttc tta ttc tca gca atg gaa tta gag aca	1296
Val Ser Gly Asp Gly Gly Phe Leu Phe Ser Ala Met Glu Leu Glu Thr	
420 425 430	
gca gtt cga cta aaa gca cca att gta cac att gta tgg aac gac agc	1344
Ala Val Arg Leu Lys Ala Pro Ile Val His Ile Val Trp Asn Asp Ser	
435 440 445	
aca tat gac atg gtg cat ttc cag caa ttg aaa aaa tat aac cgt aca	1392
Thr Tyr Asp Met Val His Phe Gln Gln Leu Lys Lys Tyr Asn Arg Thr	
450 455 460	
tct gcg gtc gat ttc gga aat atc gat atc gtg aaa tat gcg gaa agc	1440
Ser Ala Val Asp Phe Gly Asn Ile Asp Ile Val Lys Tyr Ala Glu Ser	
465 470 475 480	
ttc gga gca act gcg ttg cgc gta gaa tca cca gac cag ctg gca gat	1488
Phe Gly Ala Thr Ala Leu Arg Val Glu Ser Pro Asp Gln Leu Ala Asp	
485 490 495	
gtt ctg cgt caa ggc atg aac gct gaa ggt cct gtc atc atc gat gtc	1536
Val Leu Arg Gln Gly Met Asn Ala Glu Gly Pro Val Ile Ile Asp Val	
500 505 510	
ccg gtt gac tac agt gat aac att aat tta gca agt gac aag ctt ccg	1584
Pro Val Asp Tyr Ser Asp Asn Ile Asn Leu Ala Ser Asp Lys Leu Pro	
515 520 525	
aaa gaa ttc ggg gaa ctc atg aaa acg aaa gct ctc tag	1623
Lys Glu Phe Gly Glu Leu Met Lys Thr Lys Ala Leu	
530 535 540	

&lt;210&gt; 87

&lt;211&gt; 540

&lt;212&gt; PRT

&lt;213&gt; Bacillus subtilis

&lt;400&gt; 87

Met Tyr Leu Ala Phe Gln Val Gln Lys Leu Met Arg Tyr Leu Thr Leu
1 5 10 15

Tyr Lys Ile Lys Asp Leu Lys Leu Ser Leu Pro Gly Thr Asn Lys Thr
20 25 30

- 133 -

Gln Gln Phe Met Ala Gln Ala Val Gly Arg Leu Thr Gly Lys Pro Gly  
 35 40 45  
 Val Val Leu Val Thr Ser Gly Pro Gly Ala Ser Asn Leu Ala Thr Gly  
 50 55 60  
 Leu Leu Thr Ala Asn Thr Glu Gly Asp Pro Val Val Ala Leu Ala Gly  
 65 70 75 80  
 Asn Val Ile Arg Ala Tyr Arg Leu Lys Arg Thr His Gln Ser Leu Asp  
 85 90 95  
 Asn Ala Ala Leu Phe Gln Pro Ile Thr Lys Tyr Ser Val Glu Val Gln  
 100 105 110  
 Asp Val Lys Asn Ile Pro Glu Ala Val Thr Asn Ala Phe Arg Ile Ala  
 115 120 125  
 Ser Ala Gly Gln Ala Gly Ala Ala Phe Val Ser Phe Pro Gln Asp Val  
 130 135 140  
 Val Asn Glu Val Thr Asn Thr Lys Asn Val Arg Ala Val Ala Ala Pro  
 145 150 155 160  
 Lys Leu Gly Pro Ala Ala Asp Asp Ala Ile Ser Ala Ala Ile Ala Lys  
 165 170 175  
 Ile Gln Thr Ala Lys Leu Pro Val Val Leu Val Gly Met Lys Gly Gly  
 180 185 190  
 Arg Pro Glu Ala Ile Lys Ala Val Arg Lys Leu Leu Lys Lys Val Gln  
 195 200 205  
 Leu Pro Phe Val Glu Thr Tyr Gln Ala Ala Gly Thr Leu Ser Arg Asp  
 210 215 220  
 Leu Glu Asp Gln Tyr Phe Gly Arg Ile Gly Leu Phe Arg Asn Gln Pro  
 225 230 235 240  
 Gly Asp Leu Leu Leu Glu Gln Ala Asp Val Val Leu Thr Ile Gly Tyr  
 245 250 255  
 Asp Pro Ile Glu Tyr Asp Pro Lys Phe Trp Asn Ile Asn Gly Asp Arg  
 260 265 270  
 Thr Ile Ile His Leu Asp Glu Ile Ile Ala Asp Ile Asp His Ala Tyr  
 275 280 285  
 Gln Pro Asp Leu Glu Leu Ile Gly Asp Ile Pro Ser Thr Ile Asn His  
 290 295 300  
 Ile Glu His Asp Ala Val Lys Val Glu Phe Ala Glu Arg Glu Gln Lys  
 305 310 315 320  
 Ile Leu Ser Asp Leu Lys Gln Tyr Met His Glu Gly Glu Gln Val Pro  
 325 330 335  
 Ala Asp Trp Lys Ser Asp Arg Ala His Pro Leu Glu Ile Val Lys Glu

- 134 -

			340					345					350			
Leu	Arg	Asn	Ala	Val	Asp	Asp	His	Val	Thr	Val	Thr	Cys	Asp	Ile	Gly	
		355					360					365				
Ser	His	Ser	Ile	Trp	Met	Ser	Arg	Tyr	Phe	Arg	Ser	Tyr	Glu	Pro	Leu	
	370					375					380					
Thr	Leu	Met	Ile	Ser	Asn	Gly	Met	Gln	Thr	Leu	Gly	Val	Ala	Leu	Pro	
385					390					395					400	
Trp	Ala	Ile	Gly	Ala	Ser	Leu	Val	Lys	Pro	Gly	Glu	Lys	Val	Val	Ser	
				405					410					415		
Val	Ser	Gly	Asp	Gly	Gly	Phe	Leu	Phe	Ser	Ala	Met	Glu	Leu	Glu	Thr	
			420					425					430			
Ala	Val	Arg	Leu	Lys	Ala	Pro	Ile	Val	His	Ile	Val	Trp	Asn	Asp	Ser	
		435					440					445				
Thr	Tyr	Asp	Met	Val	His	Phe	Gln	Gln	Leu	Lys	Lys	Tyr	Asn	Arg	Thr	
	450					455					460					
Ser	Ala	Val	Asp	Phe	Gly	Asn	Ile	Asp	Ile	Val	Lys	Tyr	Ala	Glu	Ser	
465					470					475					480	
Phe	Gly	Ala	Thr	Ala	Leu	Arg	Val	Glu	Ser	Pro	Asp	Gln	Leu	Ala	Asp	
				485					490					495		
Val	Leu	Arg	Gln	Gly	Met	Asn	Ala	Glu	Gly	Pro	Val	Ile	Ile	Asp	Val	
			500					505					510			
Pro	Val	Asp	Tyr	Ser	Asp	Asn	Ile	Asn	Leu	Ala	Ser	Asp	Lys	Leu	Pro	
		515					520					525				
Lys	Glu	Phe	Gly	Glu	Leu	Met	Lys	Thr	Lys	Ala	Leu					
	530					535					540					

&lt;210&gt; 88

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: ribosome binding site

&lt;220&gt;

&lt;223&gt; All occurrences of n indicate any nucleotide

&lt;400&gt; 88

agaaaggagg tgannnnnnn atg

23

&lt;210&gt; 89

&lt;211&gt; 7

&lt;212&gt; PRT

- 135 -

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: PanC  
C terminus

&lt;400&gt; 89

Ile Arg Glu Met Glu Arg Ile  
1 5

&lt;210&gt; 90

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: PanC  
C terminus

&lt;400&gt; 90

Ile Arg Glu Arg Arg  
1 5

&lt;210&gt; 91

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: PanC  
C terminus

&lt;400&gt; 91

Ile Arg Arg Lys Glu Val Asn  
1 5

&lt;210&gt; 92

&lt;211&gt; 6688

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
pAN336 plasmid

&lt;400&gt; 92

tgcgcccgtc cagggcgctt ccattcgcca ttcaggctgc gcaactgttg ggaagggcga 60  
tcgggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 120  
ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagttaa 180  
ttgtaatacg actcactata gggcggaattg ggcccgcacgt cgcattgcacc aggccttctca 240  
ggcgctgact tagaaaacct cttgaatgaa gctgcgcttg tagcggctcg tcaaaacaag 300

aaaaaaatcg atgcgcgtga tattgacgaa gcgacggacc gtgtaattgc cggacccgct 360  
aagaagagcc gcgttatctc caagaaagaa cgcaatatcg tggcttatca cgaaggcggga 420  
cacaccgtta tcggtctcgt tttagatgag gcagatatgg ttcataaagt aacgattgtt 480  
cctcggggcc aggctggcgg ttatgctgtt atgctgccaa gagaagaccg ttatttccaa 540  
acaaagccgg agctgcttga taaaattgtc ggcctcttgg gcggacgtgt tgctgaagag 600  
attatcttcg gtgaagtcag cacaggggcg cacaatgact tccagcgtgc gacgaatatt 660  
gcaagacgaa tggttacaga attcggtatg tcagaaaaac tgggaccgtt gcaatttgga 720  
cagtctcagg gcggtcagggt attcttaggc cgtgatttca acaacgaaca gaactacagt 780  
gatcaaatcg cttacgaaat tgatcaggaa attcagcgca tcatcaaaga atgttatgag 840  
cgtgcgaaac aaatcctgac tgaaaatcgt gacaagcttg aattgattgc ccaaacgctt 900  
ctgaaagttg aaacgcttga cgctgaacaa atcaaacacc ttatcgatca tggaacatta 960  
cctgagcgta atttctcaga tgatgaaaag aacgatgatg tgaaagtaaa cattctgaca 1020  
aaaacagaag aaaagaaaga cgatacgaag gagtaattcg ctttctttct aaaaaaactg 1080  
ccggctgacg ctggcagttt ttttatgtaa atgattggct cagctgcggc ttttacaatc 1140  
atccaattct ggtatcgatt tgtttacaaa tgagccgctg atcgtgtatg gtattgtaga 1200  
atgtttgtaa aaagtaaagt agagaaacta ttcaaaagtg gtgatagagg ttgttactgg 1260  
ttatcgatgt ggggaacacc ctgcagctcg agtgaaatac cgcacagatg cgtgaaggaga 1320  
aaataccgca tcaggcgata aaccagcgca accatttgag gtgataggta agattatacc 1380  
gaggtatgaa aacgagaatt ggacctttac agaattactc tatgaagcgc catattttaa 1440  
aagctaccaa gacgaagagg atgaagagga tgaggaggca gattgccttg aatatattga 1500  
caatactgat aagataatat atcttttata tagaagatat cgccgtatgt aaggatttca 1560  
gggggcaagg cataggcagc gcgcttatca atatatctat agaatgggca aagcataaaa 1620  
acttgcatgg actaatgctt gaaaccagc acaataacct tatagcttgt aaattctatc 1680  
ataattgtgg tttcaaaatc ggctccgctg atactatgtt atacgccaac tttcaaaaca 1740  
actttgaaaa agctgttttc tggattttaa ggttttagaa tgcaaggaac agtgaattgg 1800  
agttcgtctt gttataatta gcttcttggg gtatctttaa atactgtaga aaagaggaag 1860  
gaaataataa atggctaaaa tgagaatatc accggaattg aaaaaactga tcgaaaaata 1920  
ccgctgcgta aaagatacgg aaggaatgtc tcctgctaag gtatataagc tgggtgggaga 1980  
aatgaaaac ctatatttaa aatgacgga cagccggtat aaagggacca cctatgatgt 2040

- 137 -

ggaacgggaa aaggacatga tgctatggct ggaaggaaaag ctgcctgttc caaaggtcct 2100  
 gcactttgaa cggcatgatg gctggagcaa tctgctcatg agtgaggccg atggcgctcct 2160  
 ttgctcggaa gagtatgaag atgaacaaaag ccctgaaaag attatcgagc tgtatgcgga 2220  
 gtgcatcagg ctctttcact ccatcgacat atcggattgt ccctatacga atagcttaga 2280  
 cagccgctta gccgaattgg attacttact gaataacgat ctggccgatg tggattgcga 2340  
 aaactgggaa gaagacactc catttaaaga tccgcgcgag ctgtatgatt ttttaaagac 2400  
 ggaaaagccc gaagaggaac ttgtcttttc ccacggcgac ctgggagaca gcaacatctt 2460  
 tgtgaaagat ggcaaagtaa gtggctttat tgatcttggg agaagcggca gggcggacaa 2520  
 gtggtatgac attgccttct gcgtccggtc gatcaggagg gatatcgggg aagaacagta 2580  
 tgtcgagcta ttttttgact tactygggat caagcctgat tgggagaaaa taaaatatta 2640  
 tattttactg gatgaattgt tttagtacct agatttagat gtctaaaaag ctttaactac 2700  
 aagcttttta gacatctaatt cttttctgaa gtacatccgc aactgtccat actctgatgt 2760  
 tttatatctt ttctaaaagt tcgctagata ggggtcccga gcgcctacga ggaatttgta 2820  
 tcgccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcgggtc ggtcgaactg 2880  
 caggcaaaac aggaccaag gtcattgcga caggaggcct ggccgcgctc attgcgaacy 2940  
 aatcagattg tatagacatc gttgatccat tcttaaccct aaaagggctg gaattgattt 3000  
 atgaaagaaa ccgcgtagga agtgtatagg aggttttagta atggattatt tagtaaaagc 3060  
 acttgcgtat gacggaaaaag ttcgggctta tgcagcgaga acgactgata tggtaaataga 3120  
 ggggcagaga cgccatggta cgtggccgac agcatccgct gcactaggcc gtacaatgac 3180  
 agcttcactt atgctcggcg ctatgctgaa gggcgatgat aagctgaccg tgaaaatcga 3240  
 gggcggagggt ccgatcggag ctattgtagc tgatgccaat gccaaaggag aagtcagagc 3300  
 ctatgtctct aacccgcaag ttcattttga tttaaatgaa caaggtaagc ttgatgtcag 3360  
 acgtgcggtt ggaacaaaac gaacgttaag tgtcgtaaaa gatttaggtt tgcgcgagtt 3420  
 cttcacagga caagtagaaa tcgtttcagg agaattagga gatgatttta cttactatct 3480  
 tgtgtcatct gagcagggtc cttcatcagt gggcgtaggt gtgctcgtaa atcctgacaa 3540  
 taccattctt gcggcagggg gctttattat tcagctgatg ccgggaacag atgatgaaac 3600  
 aatcacaaaa attgaacagc gtctatctca agtagagccg atttctaagc tcatcaaaaa 3660  
 agggctgaca ccagaagaaa ttttagaaga agtcctaggc gagaaacctg agattttgga 3720  
 aacgatgcct gtcagattcc attgcccttg ttcaaaaagaa cggttcgaaa cagccatttt 3780



- 138 -

aggactaggc aaaaaagaaa ttcaagatat gatagaagaa gatggacaag ccgaagcagt 3840  
atgccatttt tgtaatgaaa agtacttatt tacaaaagaa gagctggaag ggcttcgtga 3900  
ccaaactacc cgctaagctc tttagcgggt ttttaatttg agaaaagggg ctgaaagcag 3960  
gtttgaaatc aagaacaatc tggacgcgtt ggatgcatag cttgagtatt ctatagtgtc 4020  
acctaaatag cttggcgtaa tcatggtcac agctgtttcc tgtgtgaaat tgttatccgc 4080  
tcacaattcc acacaacata cgagccggaa gcataaagt taaagcctgg ggtgcctaata 4140  
gagtgaagct actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 4200  
tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg 4260  
ggcgtctctc cgcttcctcg ctactgact cgctgcgtc ggctcgttcgg ctgcggcgag 4320  
cggtatcagc tcaactcaag gcggttaatac gggtatccac agaactcaggg gataacgcag 4380  
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 4440  
tggcgttttt cgataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 4500  
agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc 4560  
tcgtgcgtc tctgttccg accctgcgc ttaccggata cctgtccgcc tttctccctt 4620  
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 4680  
ttcgtccaa gctgggctgt gtgcacgaac ccccggttca gcccgaccgc tgcgccttat 4740  
ccggtaaacta tcgtcttgag tccaaccgg taagacacga cttatcgcca ctggcagcag 4800  
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 4860  
gggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc 4920  
cagttacctt cgaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta 4980  
gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag 5040  
atcctttgat cttttctac gggctctgac ctacgtggaa cgaaaactca cgtaagggga 5100  
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa 5160  
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa 5220  
tcagtgaaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gctgactcc 5280  
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga 5340  
taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa 5400  
gggcccagcg cagaagtggg cctgcaactt tatccgcctc catccagtct attaatgtt 5460  
gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacggt gttggcattg 5520

- 139 -

ctacaggcat cgtggtgtca cgctcgctgt ttggtatggc ttcattcagc tccggttccc 5580  
 aacgatcaag gcgagttaca tgatcccca tggtgtgcaa aaaagcggtt agtccttcg 5640  
 gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 5700  
 cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 5760  
 actcaaccaa gtcattctga gaataccgcg cccggcgacc gagttgctct tgcccgcgct 5820  
 caatacggga taatagtgtg tgacatagca gaactttaaa agtgctcatc attggaatac 5880  
 gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac 5940  
 ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcggt tctgggtgag 6000  
 caaaaacagg aaggcaaaat gccgcaaaaa aggaataag ggcgacacgg aaatggtgaa 6060  
 tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 6120  
 gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 6180  
 cccgaaaagt gccacctgta tgcggtgtga aataccgcac agatgcgtaa ggagaaaata 6240  
 ccgcatacagg cgaaattgta aacgttaata ttttggtaaa attcgcgtta aatatttgtt 6300  
 aaatcagctc attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaaag 6360  
 aatagaccga gatagggttg agtggtgttc cagtttggaa caagagtcca ctattaaaga 6420  
 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 6480  
 aaccatcacc caaatcaagt tttttgcggt cgagggtgccg taaagctcta aatcggaacc 6540  
 ctaaaggagg ccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg 6600  
 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc 6660  
 gcgtaaccac cacacccgcc gcgcttaa 6688

&lt;210&gt; 93

&lt;211&gt; 8503

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Recombinant  
 pAN004 plasmid

&lt;400&gt; 93

gaattttgcg gccgcttcga aagctgtaat ataaaaacct tcttcaacta acggggcagg 60  
 ttagtgacat tagaaaaccg actgtaaaaa gtacagtcgg cattatctca tattataaaa 120  
 gccagtcatt aggcctatct gacaattcct gaatagagtt cataaacaat cctgcatgat 180  
 aaccatcaca aacagaatga tgtacctgta aagatagcgg taaatatatt gaattacctt 240

tattaatgaa ttttcctgct gtaataatgg gtagaaggta attactatta ttattgatat 300  
ttaagttaaa ccagtaaat gaagtccatg gaataataga aagagaaaaa gcattttcag 360  
gtataggtgt tttgggaaac aatttccccg aaccattata tttctctaca tcagaaaggt 420  
ataaatcata aaactctttg aagtcattct ttacaggagt ccaaatacca gagaatgttt 480  
tagatacacc atcaaaaatt gtataaagtg gctctaactt atcccaataa cctaactctc 540  
cgtcgctatt gtaaccagtt ctaaaagctg tatttgagtt tatcaccctt gtcactaaga 600  
aaataaatgc agggtaaaat ttatatcctt cttgttttat gtttcggtat aaaactactaa 660  
tatcaatttc tgtggttata ctaaaagctg tttgttggtt caaataatga ttaaataatct 720  
cttttctctt ccaattgtct aaatcaattt tattaaggtt catttgatat gcctcctaaa 780  
tttttatcta aagtgaattt aggaggctta cttgtctgct ttcttcatta gaatcaatcc 840  
ttttttaaaa gtcaatatta ctgtaacata aatatatatt taaaaatat ccacttttat 900  
ccaattttcg tttgttgaac taatgggtgc tttagttgaa gaataaagac cacattaataa 960  
aatgtggtct tttgtgtttt tttaaaggat ttgagcgtag cgaaaaatcc ttttctttct 1020  
tatcttgata ataagggtaa ctattgaatt cggtagcaag agtttgtaga aacgcaaaaa 1080  
ggccatccgt caggatggcc ttctgcttaa tttgatgcct ggcagtttat ggcgggcgtc 1140  
ctgcccgcc caactccgggc cgttgcttcg caacgttcaa atccgctccc ggcggatttg 1200  
tcctactcag gagagcgttc accgacaaac aacagataaa acgaaaggcc cagtctttcg 1260  
actgagcctt tcgttttatt tgatgcctgg cagttcccta ctctcgcatg gggagacccc 1320  
aactaccat cggcgctacg gcgtttcact tctgagttcg gcatggggtc aggtgggacc 1380  
accgcgtac tgcgcgagg caaattctgt tttatcagac cgcttctgcg ttctgattta 1440  
atctgtatca ggctgaaaat cttctctcat ccgccaaaac aggatcctac ggaaatggag 1500  
cggcaaaaacc gttttactct caaaatctta aaagaaaacc cccgataaag ggggcttttc 1560  
ttctacaaaa ttgtacgggc tggttcgttc ccagcattt gttcaatttt gttttgatca 1620  
ttcagaacag ccaacttcgg ctcatggctt gccgttctt gatcagacat cttttgtag 1680  
gaaataataa tgacctatc tccttctgc acaaggcgtg cggctgcacc gtttaagcat 1740  
atgacgcgc tttcccgtt accaggaata atatacgtt caagacgtgc tccattatta 1800  
ttattcacia tttgtacttt ttcattagga agcattccca cagcatcaat gagatcttca 1860  
tcaattgtaa tgcttccac atagttcagg tttgcttccg taacagttgc cctgtgaagt 1920  
ttgccgctca tcattgttcg atacatatta tattctctcc atttctcgaa tatcaataat 1980

gatattatct attaaacgcg cttttgaaaa agcaactgca acagcgagaa tcattctttcc 2040  
agcaatttca ttcacaggct cgagttccgg ataggaataa agctctacat agtctatggc 2100  
tccgctagtc gtttcaatga tatcttttgc agctttttatc accgcttcag gatctctttc 2160  
accggcttgg acaagttccg cacttgtttg aagggcccga tacagcttag gcgcttcttt 2220  
tctttctctca gctgttaagt atacattgcg agagcttttg gctaagccgt ctctctctct 2280  
gaccgtatcg acaggaacca attcaat'atc catgaagaag tcgctgatta acccatcaac 2340  
aacagctacc tgctgcgcat cttttaaacg gaaataggca cgagtcggct tgactagatt 2400  
gaaaagcttc gtcagtacga tcgcgacccc gtcaaaatgt ccttctcttg agcggccgca 2460  
taacacgtct gtgcgtcttt ctacatgaat cgtgacattc ttttcaccgg gatacatatc 2520  
atgagcatct ggcgtaaaaa gaatatcgac tccggcgttt tctgcaagag ctgcatcccg 2580  
ctcaatatcg cgcggatatg cttcaaaatc ttcattaggc ccgaattgtg caggattcac 2640  
aaaaatactc ataataacgg cgtcgttttc ttgtcttgct ttgtctgcta aggttaaagt 2700  
cccctcatgc agaaacccca tcgtcggaac aaatccgatt gacttgccct ctgaatggta 2760  
ttgttttatg gcttctttca gctgtgaaat atcagtaatc tgtctcatct tattttcccc 2820  
cgtacaagcc gtcaagcact gtctggttca tttgaaagga atgcttttgt tcagggaag 2880  
cacgatgtct tacatcctga acatatccgc tgattgctgt ttcgatgggt tcatcaatgc 2940  
gcgtatattg ctttacaagt ttaggtgttc tctcaacacc gtggccgata atatcatgat 3000  
aaacgagAAC ttgtccgtcc gctttcacac cagccccgat tccaatgacc ggtatgctta 3060  
gcgtctcggc aattttggct gtgagttctg ccggcacaca ttccagcaca agcatcatag 3120  
ctctgcttc ttcgcatttt atactgtct ctattaattt tttggcgctt tgttcgtctt 3180  
tgccctgtac tttatagccg ccagtaacgc cgactgactg cgggtgtcaaa cctaagtgc 3240  
tgactactgg aatgcctcca agcgtcaatg cgcgaatgga ttcaaacacg ccttctccgc 3300  
cctcaagctt cagtgcgtca gctccgcttt cctgaacgat agccgctgca tttttcagcg 3360  
tatcttctct agacaggtga taagacataa acggcatatc tgtcacaata aaggatttcg 3420  
gcgcacccct tttaacggct tttgtatgat ggatcatgtc cgcaactgtc acaccgacag 3480  
ttgaatcaag gccgaggacg accattccaa gtgaatcacc gactaaaatc atgtcaactc 3540  
ccgcttgctc agcaagttta gctgcgggat aatcataagc ggtcagcatg acaatcggtt 3600  
cttcagactc cttcattttt agaaaatcca gttttgtttt catgttttct cctccttcta 3660  
gagcgtctcg ctgttggtta gattattata ccacaccttg tagataaagt caacaacttt 3720

ttgcaaaatt tttcaggaat ttttagcagag gttgttctgg atgtagaaca aaacatcttt 3780  
ccgctcttgt gctgttagga tatctttctt ggaagctagg taggcctcga gttatggcag 3840  
ttgggttaaaa ggaaacaaaa agaccgtttt cacacaaaac ggtctttttc gatttctttt 3900  
tacagtcaca gccacttttg caaaaaccgg acagcttcat gccttataac tgctgtttcg 3960  
gtcgacaagc ttcgcgaagc ggccgcaaaa ttcactggcc gtcgttttac aacgtcgtga 4020  
ctgggaaaaac cctggcgta cccaacttaa tcgccttgca gcacatcccc ctttcgccag 4080  
ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc gcagcctgaa 4140  
tggcgaatgg cgctgatgc ggtattttct ccttacgcat ctgtgcggtta tttcacaccg 4200  
catatgggtgc actctcagta caatctgctc tgatgccgca tagttaagcc agccccgaca 4260  
cccgccaaca cccgctgact atgcttgtaa accgttttgt gaaaaaattt ttaaaataaa 4320  
aaaggggacc tctaggggcc ccaattaatt agtaataata tctattaaag gtcattcaaa 4380  
aggtcatcca ccgatcagc ttagtaaagc cctcgctaga ttttaatgcg gatgttgcca 4440  
ttacttcgcc aactattgcg ataacaagaa aaagccagcc tttcatgata tatctcccaa 4500  
tttgtgtagg gcttattatg cacgcttaaa aataataaaa gcagacttga cctgatagtt 4560  
tggtctgtgag caattatgtg cttagtgcac ctaacgcttg agttaagccg cgccgcgaag 4620  
cggcgtcgyc ttgaacgaat tgtagacat tatttgccga ctaccttggt gatctgcct 4680  
ttcacgtagt ggacaaattc ttccaactga tctgcgcgcy aggccaagcg atcttcttct 4740  
tgtccaagat aagcctgtct agcttcaagt atgacgggct gatactgggc cggcaggcgc 4800  
tccattgccc agtcggcagc gacatccttc ggcgcgattt tgccggttac tgcgctgtac 4860  
caaatgcggg acaacgtaag cactacattt cgctcatcgc cagcccagtc gggcggcgag 4920  
ttccatagcg ttaaggtttc atttagcgcc tcaaatagat cctgttcagg aaccggatca 4980  
aagagttcct ccgcccgtgg acctaccaag gcaacgctat gttctcttgc ttttgtcagc 5040  
aagatagcca gatcaatgtc gatcgtggct ggctcgaaga tacctgcaag aatgtcattg 5100  
cgctgccatt ctccaaattg cagttcgcgc ttagctggat aacgccagc aatgatgtcg 5160  
tcgtgcacaa caatggtgac ttctacagcg cggagaatct cgctctctcc aggggaagcc 5220  
gaagtttcca aaaggtcggt gatcaaagct cgccgcggtt tttcatcaag ccttacggtc 5280  
accgtaacca gcaaatcaat atcactgtgt ggcttcaggc cgccatccac tgcggagccg 5340  
tacaaatgta cggccagcaa cgtcggttcg agatggcgct cgatgacgcc aactacctct 5400  
gatagttgag tcgatacttc ggcgatcacc gcttccctca tgatgtttaa ctttgtttta 5460

gggcgactgc cctgctgcgt aacatcgttg ctgctccata acatcaaaca tcgaccacg 5520  
gcgtaacgcg cttgctgctt ggatgcccg ggcatagact gtaccccaaa aaaacagtca 5580  
taacaagcca tgaaaaccgc cactgcgcg ttaccaccgc tgcgttcggt caaggttctg 5640  
gaccagttgc gtgagcgcgt acgctacttg cattacagct tacgaaccga acaggcttat 5700  
gtccactggg ttctgtgcct catccgtttc cacggtgtgc gtcaccggc aaccttgggc 5760  
agcagcgaag tcgaggcatt tctgtcctgg ctggcgaacg agcgaagggt ttcggtctcc 5820  
acgcacgctc aggcatggc ggcttgctg ttcttctacg gcaagggtgt gtgcacggat 5880  
ctgccctggc ttcaggagat cggaagacct cggccgtcgc ggcgcttgcc ggtggtgctg 5940  
accccggtatg aagtggttcg catcctcggg tttctggaag gcgagcatcg tttgttcgcc 6000  
cagcttctgt atggaacggg catgcggatc agtgagggtt tgcaactgcg ggtcaaggat 6060  
ctggatttcg atcacggcac gatcatcgtg cgggagggca agggctccaa ggatcgggcc 6120  
ttgatgttac ccgagagctt ggcaccacgc ctgcgcgagc aggggaattg atccggtgga 6180  
tgaccttttg aatgacctt aatagattat attactaatt aattggggac cctagaggtc 6240  
ccctttttta ttttaaaaat tttttcaca aacggtttac aagcataacg ggttttgctg 6300  
cccgcaaacg ggctgttctg gtgttgctag tttgttatca gaatcgcaga tccgggttca 6360  
ggtttgcccg ctgaaagcgc tatttcttcc agaattgcc aattttttc cccacgggag 6420  
gcgtcactgg ctcccggtgt gtcggcagct ttgattcgat aagcagcatc gcctgtttca 6480  
ggctgtctat gtgtgactgt tgagctgtaa caagttgtct caggtgttca atttcatgtt 6540  
ctagttgctt tgttttactg gtttcacctg ttctattagg tgttacatgc tgttcatctg 6600  
ttacattgtc gatctgttca tggatgaacag ctttaaattgc accaaaaact cgtaaaagct 6660  
ctgatgtatc tatctttttt acaccgtttt catctgtgca tatggacagt tttccctttg 6720  
atatctaacg gtgaacagtt gttctacttt tgtttgtag tcttgatgct tctactgatg 6780  
atacaagagc cataagaacc tcagatcctt ccgtatttag ccagtatgtt ctctagtgtg 6840  
gttcgttggt tttgcgtgag ccatgagaac gaaccattga gatcatgctt actttgcatg 6900  
tactcaaaa attttgctc aaaactggtg agctgaattt ttgcagttaa agcatcgtgt 6960  
agtgtttttc ttagtccgtt acgtaggtag gaatctgatg taatggttgt tggatatttg 7020  
tcaccattca tttttatctg gttgttctca agttcgggta cgagatccat ttgtctatct 7080  
agttcaactt ggaaaatcaa cgtatcagtc gggcggcctc gcttatcaac caccaatttc 7140  
atattgctgt aagtgtttaa atctttactt attggtttca aaaccattg gttaagcctt 7200

- 144 -

```

ttaaactcat ggtagtatt ttcaagcatt aacatgaact taaattcatc aaggctaatac 7260
tctatatttg ccttggtgagt tttcttttgt gttagttctt ttaataacca ctcataaatc 7320
ctcatagagt atttggtttc aaaagactta acatgttcca gattatattt tatgaatttt 7380
tttaactgga aaagataagg caatatctct tcaactaaaa ctaattctaa tttttcgctt 7440
gagaacttgg catagtttgt ccaactggaaa atctcaaagc cttaaaccaa aggattcctg 7500
atttccacag ttctcgtcac cagctctctg gttgctttag ctaatacacc ataagcattt 7560
tccctactga tgttcatcat ctgagcgtat tgggtataag tgaacgatac cgtccgttct 7620
ttccttgtag ggttttcaat cgtgggggtg agtagtgcca cacagcataa aattagcttg 7680
gtttcatgct cgttaagtc atagcgacta atcgctagtt catttgcttt gaaaacaact 7740
aattcagaca tacatctcaa ttggtctagg tgattttaat cactatacca attgagatgg 7800
gctagtcaat gataattact agtccttttc ctttgagttg tgggtatctg taaattctgc 7860
tagacctttg ctggaaaact tgtaaattct gctagaccct ctgtaaattc cgctagacct 7920
ttgtgtgttt tttttgttta tattcaagtg gttataattt atagaataaa gaaagaataa 7980
aaaaagataa aaagaataga tcccagccct gtgtataact cactacttta gtcagttccg 8040
cagtattaca aaaggatgct gcaaacgctg tttgctcctc taaaaacag accttaaaac 8100
cctaaagggt taagtagcac cctcgcaagc tcgggcaaatt cgctgaatat tccttttgtc 8160
tcggaccatc aggcacctga gtcgctgtct ttttcgtgac attcagttcg ctgcgctcac 8220
ggctctggca gtgaatggg gtaaattggca ctacaggcgc cttttatgga ttcattgcaag 8280
gaaactaccc ataatacaag aaaagcccgt cacgggcttc tcagggcgtt ttatggcggg 8340
tctgctatgt ggtgctatct gactttttgc tgttcagcag ttctgcct ctgattttcc 8400
agtctgacca cttcggatta tcccgtgaca ggctattcag actggctaata gcaccagta 8460
aggcagcggg atcatcaaca ggcttaccgc tcttactgtc aac 8503

```

&lt;210&gt; 94

&lt;211&gt; 7381

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: recombinant  
pAN006 plasmid

&lt;400&gt; 94

```

ttgcggccgc ttcgaaagct gtaatataaa aaccttcttc aactaacggg gcagggttagt 60

```

- 145 -

gacattagaa aaccgactgt aaaaagtaca gtcggcatta tctcatatta taaaagccag 120  
 tcattaggcc tatctgacaa ttcttgaata gagttcataa acaatcctgc atgataacca 180  
 tcacaaacag aatgatgtac ctgtaaagat agcggtaaatt atattgaatt acctttatta 240  
 atgaattttc ctgctgtaatt aatgggtaga aggtaattac tattattatt gatatttaag 300  
 ttaaaccag taaatgaagt ccatggaata atagaaagag aaaaagcatt ttcaggata 360  
 ggtgttttgg gaaacaattt ccccgaaacca ttatatttct ctacatcaga aaggtataaa 420  
 tcataaaaact ctttgaagtc attctttaca ggagtccaaa taccagagaa tgtttttagat 480  
 acaccatcaa aaattgtata aagtggctct aacttatccc aataacctaa ctctcgtcgc 540  
 ctattgtaac cagttctaaa agctgtattt gagtttatca cccttggtcac taagaaaata 600  
 aatgcagggt aaaatttata tccttcttgt tttatgtttc ggtataaaac actaatatca 660  
 atttctgtgg ttatactaaa agtcgtttgt tggttcaaatt aatgattaaa tatctctttt 720  
 ctcttccaat tgtctaaatc aattttatta aagttcattt gatatgcctc cttaaattttt 780  
 atctaaagtg aatttaggag gcttacttgt ctgctttctt cattagaatc aatccttttt 840  
 taaaagtcaa tattactgta acataaatat atattttaaa aatatcccac tttatccaat 900  
 tttcgtttgt tgaactaatg ggtgcttttag ttgaagaata aagaccacat taaaaaatgt 960  
 ggtcttttgt gtttttttaa aggatttgag cgtagcgaaa aatccttttc tttcttatct 1020  
 tgataataag ggtaactatt gaattcggta ccaagagttt gtagaaacgc aaaaaggcca 1080  
 tccgtcagga tggccttctg cttaatttga tgccctggcag tttatggcgg gcgtcctgcc 1140  
 cgccaccctc cgggcccgtt ctccgcaacg ttcaaaccg ctcccggcgg atttgtccta 1200  
 ctcaggagag cgttcaccga caaacaacag ataaaacgaa aggcccagtc tttcgactga 1260  
 gcctttcgtt ttatttgatg cctggcagtt ccctactctc gcatggggag accccacact 1320  
 accatcggcg ctacggcggt tcaacttctga gttcggcatg gggtcagggt ggaccaccgc 1380  
 gctactgccg ccaggcaaat tctgttttat cagaccgctt ctgcgttctg atttaactctg 1440  
 tatcaggctg aaaatcttct ctcatccgcc aaaacaggat cctacggaaa tggagcggca 1500  
 aaaccgtttt actctcaaaa tcttaaaaaga aaacccccga taaagggggc tttcttcta 1560  
 caaaattgta cgggctgggt cgttccccag catttggtca atttgtttt gatcattcag 1620  
 aacagccact ttcggctcat ggcttgccgc ttcttgatca gacatcattt tgtaggaaat 1680  
 aataatgacc ttatctcctt cctgcacaag gcgtgcccgt gcaccgttta agcatatgac 1740  
 gccgcttccc cgtttaccag gaataatata cgtttcaaga cgtgctccat tattattatt 1800



- 146 -

cacaatttgt actttttcat taggaagcat tcccacagca tcaatgagat cttcatcaat 1860  
tgtaatgctt cccacatagt tcaggtttgc ttccgtaaca gttgccctgt gaagtttgcc 1920  
gctcatcatt gttcgataca tattatattc tctccatttc tcgaatatca ataatgatat 1980  
tatctattaa acgcgctttt gaaaaagcaa ctgcaacagc gagaatcatc tttccagcaa 2040  
tttcattcac aggctcgagt tccggatagg aataaagctc tacatagtct atgggtccgc 2100  
tagtcgtttc aatgatatac tttgcagctt ttatcacgcg ttcaggatct ctttcaccgg 2160  
cttggaacag ttccgcactt gtttgaaggc ccgatacag cttaggcgct tcttttcttt 2220  
cctcagctgt taagtataca ttgcgagagc ttttgctaa gccgtcttcc tctctgaccg 2280  
tatcgacagg aaccaattca atatccatga agaagtcgct gattaacca tcaacaacag 2340  
ctacctgctg cgcatctttt aaaccgaaat aggcacgagt cggcttgact agattgaaaa 2400  
gcttcgtcag tacgatcgcg acccgtcaa aatgtccttc tcttgagcgc ccgcataaca 2460  
cgtctgtgcg tctttctaca tgaatcgtga cattcttttc accgggatac atatcatgag 2520  
catctggcgt aaaaagaata tcgactccgg cgttttctgc aagagctgca tcccgtcaa 2580  
tctcgcgcg atagtcttca aaatcttcat tagggccgaa ttgtgcagga ttcacaaaaa 2640  
tactcataat aacggcgctg ttttcttgtc ttgctttgtc tgctaagggt aaatgccct 2700  
catgcagaaa ccccatcgtc ggaacaaatc cgattgactt gccctctgaa tggatttgtt 2760  
ttatggcttc tttcagctgt gaaatatcag taatctgtct catcttattt tccccgtac 2820  
aagccgtcaa gcactgtctg gttcatttga aaggaatgct tttgttcagg gaaagcacga 2880  
tgtcttacat cctgaacata tccgctgatt gctgtttcga tggtttcac aatgcgcgta 2940  
tattgcttta caaatttagg tggtctctca acaccgtggc cgataatatc atgataaacg 3000  
agaacttgtc cgtccgcttt cacaccagcc ccgattccaa tgaccggtat gcttagcgtc 3060  
tcggcaattt tggtgtgag ttctgccggc acacattcca gcacaagcat catagctcct 3120  
gcttcttcgc attttatact gtcttctatt aattttttgg cgctttgttc gtctttgccc 3180  
tgtactttat agccgccag tacgccgact gactgcggtg tcaaacctaa gtgactgact 3240  
actggaatgc ctccaagegt caatgcgcga atggattcaa acacgccttc tccgccctca 3300  
agcttcagtg cgtcagctcc gctttctga acgatagccg ctgcattttt cagcgtatct 3360  
tccttagaca ggtgataaga cataaacggc atatctgtca caataaagggt attcggcgca 3420  
ccccctttaa cggcttttgt atgatggatc atgtccgcaa ctgtcacacc gacagttgaa 3480  
tcaaggccga ggacgacct tccaagtga taccgacta aaatcatgtc aactcccgct 3540

- 147 -

tgttcagcaa gtttagctgc cggataatca taagcggcca gcatgacaat cggttcttca 3600  
gactccttca tttttagaaa atccagtttt gttttcatgt tttctcctcc tctagagcgt 3660  
cctgctgttg ttaagattat tataccacac cttgtagata aagtcaacaa ctttttgcaa 3720  
aatttttcag gaatttttagc agaggttggt ctggatgtag aacaaaacat ctttccgctc 3780  
ttgtgctgtt aggatatctt tcttggaagc taggtaggcc tcgagttatg gcagttgggt 3840  
aaaaggaaac aaaaagaccg ttttcacaca aaacggctct tttcgatttc tttttacagt 3900  
cacagccact tttgcaaaaa ccggacagct tcatgcctta taactgctgt ttcggctcgac 3960  
ctgcaggcat gcaagcttcg cgaagcggcc gccgacgca ggctggatgg ccttccccat 4020  
tatgattctt ctgccttccg gcggcatcgg gatgcccgcg ttgcaggcca tgctgtccag 4080  
gcaggtagat gacgaccatc agggacagct tcaaggatcg ctgcgggctc ttaccagcct 4140  
aacttcgatc actggaccgc tgategtcac ggcgatttat gccgcctcgg cgagcacatg 4200  
gaacgggttg gcatggattg taggcgcgcg cctatacctt gtctgcctcc ccgcgttgcg 4260  
tcgcggtgca tggagccggg ccacctcgac ctgaatggaa gccggcgcca cctcgctaac 4320  
ggattcacca ctccaagaat tggagccaat caattcttgc ggagaactgt gaatgcgcaa 4380  
accaaccctt ggcagaacat atccatcgcg tccgccatct ccagcagccg cacgcggcgc 4440  
atctcgggca gcgttgggtc ctggccacgg gtgcgcata tctgtctctt gtcgttgagg 4500  
acccggttag gctggcgggg ttgccttact ggtagcaga atgaatcacc gatacgcgag 4560  
cgaacgtgaa gcgactgctg ctgcaaaacg tctgcgacct gagcaacaac atgaatggtc 4620  
ttcggtttcc gtgtttcgta aagtctggaa acgcggaagt cagcgccctg caccattatg 4680  
ttccggatct gcatcgagg atgctgctgg ctaccctgtg gaacacctac atctgtatta 4740  
acgaagcgct ggcattgacc ctgagtgatt tttctctggt cccgccgcat ccataccgcc 4800  
agttgtttac cctcacaacg ttccagtaac cgggcatgtt catcatcagt aaccgtatc 4860  
gtgagcatcc tctctcgttt catcggtatc attaccccca tgaacagaaa ttccccctta 4920  
cacggaggca tcaagtgacc aaacaggaaa aaaccgccct taacatggcc cgctttatca 4980  
gaagccagac attaacgctt ctggagaaac tcaacgagct ggacgcggat gaacaggcag 5040  
acatctgtga atcgcttcac gaccacgtg atgagcttta ccgcagctgc ctgcgcggtt 5100  
tcggtgatga cggtgaaaac ctctgacaca tgcagctccc ggagacggtc acagcttgct 5160  
tgtaagcgga tgccgggagc agacaagccc gtcagggcgc gtcagcgggt gttggcgggt 5220  
gtcggggcgc agccatgacc cagtcacgta gcgatagcgg agtgtatact ggcttaacta 5280

- 148 -

tgcggcatca gagcagattg tactgagagt gcaccatatg cgggtgtgaaa taccgcacag 5340  
atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct tcctcgctca ctgactcgct 5400  
gcgctcggtc gttcggctgc ggcgagcggc atcagctcac tcaaaggcgg taatacggtt 5460  
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 5520  
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggtccgccc cccctgacga 5580  
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 5640  
ccaggcgttt ccccttgaa gctccctcgt gcgctctcct gtccgaccc tgccgcttac 5700  
cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 5760  
taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 5820  
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag 5880  
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatatgt 5940  
aggcgggtgct acagagttct tgaagtggcg gcctaactac ggctacacta gaaggacagt 6000  
atattggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 6060  
atccggcaaa caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac 6120  
ggcgagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggc ctgacgctca 6180  
gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac 6240  
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 6300  
ttgggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 6360  
tcgttcaccc atagttgcct gactccccgt cgtgtagata actacgatac gggagggcctt 6420  
accatctggc ccagtgctg caatgatacc gcgagacca cgctcaccgg ctccagattt 6480  
atcagcaata aaccagccag ccggaaggcg cgagcgcaga agtggctctg caactttatc 6540  
cgctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 6600  
tagtttgccg aacgttggtg ccattgctgc aggcacgtg gtgtcacgct cgctggttg 6660  
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatggtt 6720  
gtgcaaaaaa gcgggttagct ccttcgggtc tccgatcgtt gtcagaagta agttggccgc 6780  
agtgttatca ctcatggtta tggcagcaat gcataattct ctactgtca tgccatccgt 6840  
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 6900  
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 6960  
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 7020

- 149 -

gctggttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt 7080  
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 7140  
aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 7200  
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 7260  
acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacgtct aagaaaccat 7320  
tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtcttcaaga 7380  
a 7381

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
29 March 2001 (29.03.2001)

PCT

(10) International Publication Number  
WO 01/21772 A3

(51) International Patent Classification<sup>7</sup>: C12N 15/52,  
15/53, 15/54, 15/60, 15/75, 9/00, 9/02, 9/10, 9/12, 9/88,  
C12P 7/42, 13/02, 13/06

(21) International Application Number: PCT/US00/25993

(22) International Filing Date:  
21 September 2000 (21.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
09/400,494 21 September 1999 (21.09.1999) US  
60/210,072 7 June 2000 (07.06.2000) US  
60/221,836 28 July 2000 (28.07.2000) US  
60/227,860 24 August 2000 (24.08.2000) US

(71) Applicant: OMNIGENE BIOPRODUCTS [US/US];  
763(d) Concord Avenue, Cambridge, MA 02138 (US).

(72) Inventors: YOCUM, Rogers, R.; 4 Orchard Lane, Lex-  
ington, MA 02420 (US). PATTERSON, Thomas, A.; 89

Church Street, North Attleboro, MA 02760 (US). HER-  
MANN, Theron; 18 Chilhowie Drive, Kinnelon, NJ 07405  
(US). PERO, Janice, G.; 20 Solomon Pierce Road, Lex-  
ington, MA 02420 (US).

(74) Agents: HANLEY, Elizabeth, A. et al.; Lahive & Cock-  
field, LLP, 28 State Street, Boston, MA 02109 (US).

(81) Designated States (*national*): AU, BR, CA, CN, CZ, HU,  
ID, IL, IN, JP, KR, LT, LV, MX, NO, PL, RU, SG, SK, TR,  
UA, ZA.

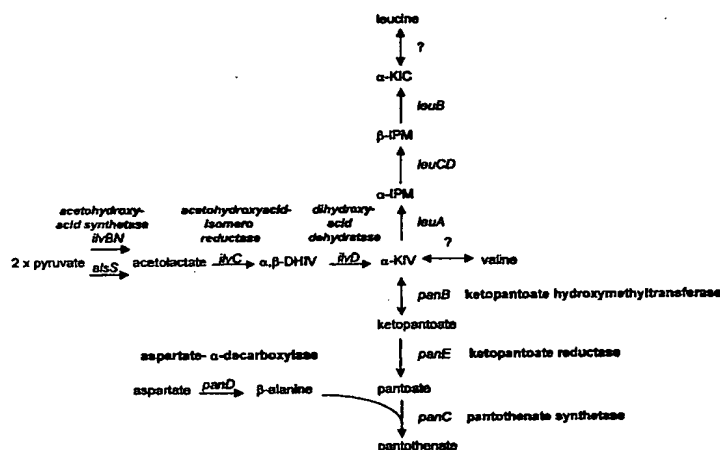
(84) Designated States (*regional*): European patent (AT, BE,  
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,  
NL, PT, SE).

Published:  
— with international search report

(88) Date of publication of the international search report:  
7 March 2002

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND MICROORGANISMS FOR PRODUCTION OF PANTO-COMPOUNDS



(57) Abstract: The present invention features methods of producing panto-compounds (<i>e.g.</i>, pantothenate) using microor-  
ganisms in which the pantothenate biosynthetic pathway and/or the isoleucine-valine biosynthetic pathway and/or the coenzymeA  
biosynthetic pathway has been manipulated. Methods featuring ketopantoate reductase overexpressing microorganisms as well as  
aspartate g(a)-decarboxylase overexpressing microorganisms are provided. Methods of producing panto-compounds in a precur-  
sor-independent manner and in high yield are described. Recombinant microorganisms, vectors, isolated nucleic acid molecules,  
genes and gene products useful in practicing the above methodologies are also provided. The present invention also features a previ-  
ously microbial pantothenate kinase gene, <i>coaX</i>, as well as methods of producing panto-compounds utilizing microorganisms  
having modified pantothenate kinase activity. Recombinant microorganisms, vectors, isolated <i>coaX</i> nucleic acid molecules  
and purified CoaX proteins are featured. Also featured are methods for identifying pantothenate kinase modulators utilizing the  
recombinant microorganisms and/or purified CoaX proteins of the present invention.

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/25993

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/52 C12N15/53 C12N15/54 C12N15/60 C12N15/75  
C12N9/00 C12N9/02 C12N9/10 C12N9/12 C12N9/88  
C12P7/42 C12P13/02 C12P13/06

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EMBL, EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, EMBASE

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 590 857 A (TAKEDA CHEMICAL INDUSTRIES LTD) 6 April 1994 (1994-04-06) cited in the application	12,13, 24,26, 27,48, 51,55, 59,60, 71,76
Y	the whole document	1-6, 33-35, 54, 56-58, 62-64, 78-82
	page 14, line 1-3 --- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

10 July 2001

Date of mailing of the international search report

22.10.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

van de Kamp, M

## INTERNATIONAL SEARCH REPORT

Int lional Application No

PCT/US 00/25993

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SAHM H ET AL.: "D-Pantothenate synthesis in Corynebacterium glutamicum and use of panBC and genes encoding L-valine synthesis for D-pantothenate overproduction" APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 65, no. 5, May 1999 (1999-05), pages 1973-1979, XP002169517 cited in the application	12,13, 24,26, 27,48, 51,54, 56,59, 60,71, 76,78,79
Y	the whole document	1-6, 33-35, 54, 56-58, 62-64, 78-82
X	--- SOROKIN A ET AL.: "Sequence analysis of the Bacillus subtilis chromosome region between the serA and kdg loci cloned in a yeast artificial chromosome" MICROBIOLOGY, vol. 142, no. 8, 1996, pages 2005-2016, XP000910121 ISSN: 1350-0872	83-86, 92,93, 95, 97-100, 102,103
Y	abstract	1-6, 33-35, 54, 56-58, 62-64, 78-82
	page 2011, right-hand column, line 14-20; table 1	
X	--- DATABASE EM_PRO [Online] EMBL; ID BSYPIA, AC L47709, 23 January 1996 (1996-01-23) HENNER D ET AL.: "Bacillus subtilis (clone YAC15-6B) ypiABF genes, qcrABC genes, ypjABCDEFGHI genes, birA gene, panBCD genes, dinG gene, ypmB gene, aspB gene, asnS gene, dnaD gene, nth gene and ypoC gene, complete cds." XP002171539 page 5, line 43-60 page 10	83-86, 98-100, 102,103
A	--- BAIGORI M ET AL.: "Isolation and characterization of Bacillus subtilis mutants blocked in the syntehsis of pantothenic acid" JOURNAL OF BACTERIOLOGY, vol. 173, no. 13, July 1991 (1991-07), pages 4240-4242, XP001002216 abstract	99,100, 102,103
	--- -/--	

# INTERNATIONAL SEARCH REPORT

In International Application No  
PCT/US 00/25993

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 18954 A (MONSANTO CO) 7 May 1998 (1998-05-07) the whole document ---	
A	EP 0 224 294 A (GIST BROCADES NV) 3 June 1987 (1987-06-03) the whole document ---	
P,X	EP 1 006 192 A (DEGUSSA) 7 June 2000 (2000-06-07)  examples 2-5 ---	12,13, 24,26, 27,48, 51, 54-56, 59,60, 71,76, 78,79
P,X	EP 1 006 189 A (DEGUSSA ;KERNFORSCHUNGSANLAGE JUELICH (DE)) 7 June 2000 (2000-06-07)  examples 1,7,9 -----	12,13, 24,26, 27,48, 51,54, 56,59, 60,71, 76,78,79



# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 00/25993

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-6,12,13,24,26-28,33-35,48,51,54-64,71,76,78-86,92,93,95,97,100,102,103  
(all Partially)

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-6,12,13,24,26-28,33-35,48,51,54-64,71,76,78-86,  
92,93,95,97-100,102,103 (all partially)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), pantoate, or ketopantoate, e.g., in a manner independent of precursor feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) which overexpresses a ketopantoate hydroxymethyltransferase-encoding gene, e.g. the *panB* gene, e.g., from *Bacillus*, under conditions such that said panto-compound is produced, and possibly further recovering the compound. A recombinant microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus* (e.g., *Bacillus subtilis*), *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*)) which overexpresses a *Bacillus* (*subtilis*) ketopantoate hydroxymethyltransferase-encoding gene. A recombinant vector encoding a *Bacillus* (*subtilis*) ketopantoate hydroxymethyltransferase-encoding gene operably linked to regulatory sequences, e.g., comprising a nucleic acid sequence according to SEQ ID NO:23 or part of SEQ ID NO:59. An isolated nucleic acid molecule encoding a *Bacillus* (*subtilis*) ketopantoate hydroxymethyltransferase, and said isolated ketopantoate hydroxymethyltransferase polypeptide.

2. Claims: 1-6,12,13,24,26-28,33-35,48,51,54-64,71,76,78-86,  
92,93,95,97-100,102,103 (all partially)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), pantoate, or ketopantoate, e.g., in a manner independent of precursor feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) which overexpresses a pantothenate synthetase-encoding gene, e.g. the *panC* gene, e.g., from *Bacillus*, under conditions such that said panto-compound is produced, and possibly further recovering the compound. A recombinant microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus* (e.g., *Bacillus subtilis*), *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*)) which overexpresses a *Bacillus* (*subtilis*) pantothenate synthetase-encoding gene. A recombinant vector encoding a *Bacillus* (*subtilis*) pantothenate synthetase-encoding gene operably linked to regulatory sequences, e.g., comprising a nucleic acid sequence according to SEQ ID NO:25 or part of SEQ ID NO:59. An isolated nucleic acid molecule encoding a *Bacillus* (*subtilis*) pantothenate synthetase, and said isolated pantothenate synthetase polypeptide.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

3. Claims: 1-6,12-14,24,26-28,33-35,48,49,51,54-64,66,71,76,  
78-86,92,93,95,97-100,102,103 (all partially); 15,  
17,19,23,32,106,107 (both completely)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), pantoate, or ketopantoate, e.g., independent of aspartate or beta-alanine feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) which overexpresses an aspartate-alpha-decarboxylase-encoding gene, e.g., from *Bacillus*, e.g., the aspartate-alpha-decarboxylase-encoding *panD* gene from *Bacillus subtilis*, under conditions such that said panto-compound is produced, and possibly further recovering the compound. A recombinant microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus* (e.g., *Bacillus subtilis*), *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*)) which overexpresses a *Bacillus* (*subtilis*) aspartate-alpha-decarboxylase-encoding gene. A recombinant vector encoding a *Bacillus* (*subtilis*) aspartate-alpha-decarboxylase-encoding gene operably linked to regulatory sequences, e.g., comprising a nucleic acid sequence according to SEQ ID NO:27 or part of SEQ ID NO:59. An isolated nucleic acid molecule encoding a *Bacillus* (*subtilis*) aspartate-alpha-decarboxylase, and said isolated aspartate-alpha-decarboxylase polypeptide.

4. Claims: 1-6,24,26-28,33-35,48,49,51,54-64,71,76,78-87,92,  
93,95,97-100,102,103 (all partially); 7-11,65,101,  
104,105 (all completely)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), pantoate, or ketopantoate, e.g., independent of precursor feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) which overexpresses a ketopantoate reductase-encoding gene, e.g., from *Bacillus*, e.g., the ketopantoate reductase-encoding *panE1* gene from *Bacillus subtilis*, under conditions such that said panto-compound is produced, and possibly further recovering the compound. A recombinant microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus* (e.g., *Bacillus subtilis*), *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*)) which overexpresses a *Bacillus* (*subtilis*) ketopantoate reductase-encoding gene. A recombinant vector encoding a *Bacillus* (*subtilis*) ketopantoate reductase-encoding gene operably linked to regulatory sequences, e.g., comprising a nucleic acid sequence according to SEQ ID NO:29. An isolated nucleic acid molecule encoding a *Bacillus* (*subtilis*)

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

ketopantoate reductase, and said isolated ketopantoate reductase polypeptide.

5. Claims: 14,16,18,28,48,54-61,66,77-82,  
97 (all partially); 20,29 (completely)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), e.g., in a manner independent of valine or alpha-ketoisovalerate feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) having a deregulated isoleucine-valine (*ilv*) pathway, wherein the microorganism overexpresses acetohydroxyacid synthase or is transformed with a vector comprising an *ilvBN* nucleic acid sequence or an *alsS* sequence, e.g., from *Bacillus*, under conditions such that pantothenate is produced, and possibly further recovering the pantothenate. Said microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus* (e.g., *Bacillus subtilis*), *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*)), and said vector.

6. Claims: 14,16,18,28,48,54-61,66,77-82,  
97 (all partially); 21,30 (completely)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), e.g., in a manner independent of valine or alpha-ketoisovalerate feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) having a deregulated isoleucine-valine (*ilv*) pathway, wherein the microorganism overexpresses acetohydroxyacid isomeroreductase or is transformed with a vector comprising an *ilvC* nucleic acid sequence, e.g., from *Bacillus*, under conditions such that pantothenate is produced, and possibly further recovering the pantothenate. Said microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus* (e.g., *Bacillus subtilis*), *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*)), and said vector.

7. Claims: 14,16,18,28,48,54-61,66,77-82,97 (partially); 22,  
31 (completely)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), e.g., in a manner independent of valine or alpha-ketoisovalerate feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci*

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

or Streptomyces) or Gram-negative) having a deregulated isoleucine-valine (ilv) pathway, wherein the microorganism overexpresses dihydroxyacid dehydratase or is transformed with a vector comprising an ilvD nucleic acid sequence, e.g., from Bacillus, under conditions such that pantothenate is produced, and possibly further recovering the pantothenate. Said microorganism (e.g., Gram-positive (e.g., belonging to the genus Bacillus (e.g., Bacillus subtilis), Corynebacterium, Lactobacillus, Lactococci or Streptomyces)), and said vector.

## 8. Claims: 25,28,50,54-61 (all partially)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), e.g., in a manner independent of precursor feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus Bacillus, Corynebacterium, Lactobacillus, Lactococci or Streptomyces) or Gram-negative) having a mutant avtA gene under conditions such that pantothenate is produced, and possibly further recovering the pantothenate.

## 9. Claims: 25,28,50,54-61 (all partially)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), e.g., in a manner independent of precursor feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus Bacillus, Corynebacterium, Lactobacillus, Lactococci or Streptomyces) or Gram-negative) having a mutant ilvE gene under conditions such that pantothenate is produced, and possibly further recovering the pantothenate.

## 10. Claims: 25,28,50,54-61 (all partially)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), e.g., in a manner independent of precursor feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus Bacillus, Corynebacterium, Lactobacillus, Lactococci or Streptomyces) or Gram-negative) having a mutant ansB gene under conditions such that pantothenate is produced, and possibly further recovering the pantothenate.

## 11. Claims: 25,28,50,54-61 (all partially)

A method of producing pantothenate (e.g., 2 g/L up to 40 g/L at least), e.g., in a manner independent of precursor feed, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus Bacillus, Corynebacterium, Lactobacillus, Lactococci or Streptomyces) or Gram-negative)

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

having a mutant *alsD* gene under conditions such that pantothenate is produced, and possibly further recovering the pantothenate.

## 12. Claims: 36,37,54-59,61 (all partially)

A method of producing beta-alanine, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) which overexpresses an aspartate-alpha-decarboxylase-encoding gene, under conditions such that beta-alanine is produced, wherein the aspartate-alpha-decarboxylase-overexpressing microorganism has a mutation in a nucleic acid sequence encoding ketopantoate hydroxymethyltransferase (*panB*), and possibly further recovering the compound.

## 13. Claims: 36,37,54-59,61 (all partially)

A method of producing beta-alanine, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) which overexpresses an aspartate-alpha-decarboxylase-encoding gene, under conditions such that beta-alanine is produced, wherein the aspartate-alpha-decarboxylase-overexpressing microorganism has a mutation in a nucleic acid sequence encoding ketopantoate reductase (*panE*), and possibly further recovering the compound.

## 14. Claims: 36,37,54-59,61 (all partially)

A method of producing beta-alanine, comprising culturing a microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) which overexpresses an aspartate-alpha-decarboxylase-encoding gene, under conditions such that beta-alanine is produced, wherein the aspartate-alpha-decarboxylase-overexpressing microorganism has a mutation in a nucleic acid sequence encoding pantothenate synthetase (*panC*), and possibly further recovering the compound.

## 15. Claim : 38 (completely)

A method of producing beta-alanine comprising contacting a composition comprising aspartate with an isolated *Bacillus* aspartate-alpha-decarboxylase enzyme under conditions such that beta-alanine is produced.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

16. Claims: 41,44-47,51,53,54-61,69,71,72,75,78-81,  
97 (all partially); 39,43,52,67,70,74,88-91,  
108-110 (all completely)

A method for producing or for enhancing production of ketopantoate, pantoate, or pantothenate (e.g., 10, 20 or 40 g/L at least), comprising culturing a mutant microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) having a mutant pantothenate kinase-encoding *coaX* gene, under conditions such that said panto-compound is produced or that production is enhanced, and possibly further recovering the compound. A method for identifying compounds which modulate pantothenate kinase activity comprising contacting a recombinant cell expressing the *coaX* gene, possibly further comprising a mutant *coaA* gene encoding a pantothenate kinase with reduced activity, with a test compound and determining the ability of the test compound to modulate pantothenate kinase activity in said cell. A recombinant microorganism having a mutant *coaX* gene encoding a pantothenate kinase with reduced activity. A vector comprising a mutant *coaX* gene encoding a pantothenate kinase with reduced activity, possibly further comprising regulatory sequences. A recombinant microorganism comprising a vector comprising an isolated *coaX* gene (e.g., from *Bacillus (subtilis)*), and said vector, possibly further comprising regulatory sequences, e.g., a constitutively active promoter. A recombinant microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus* (e.g., *Bacillus subtilis*), *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*)) that overproduces a panto-compound having a mutation in a *coaX* gene that results in a reduced level of pantothenate kinase activity, resulting in a decrease in the capacity of the microorganism to synthesize coenzyme A. An isolated nucleic acid molecule comprising a (mutant) *coaX* gene, and an isolated pantothenate kinase protein encoded by a *coaX* gene.

17. Claims: 41,44-47,51,53-61,69,71,72,75,78-81,  
97 (all partially); 40,42,68,73 (all completely)

A method for producing or for enhancing production of a panto-compound, e.g., ketopantoate, pantoate or pantothenate (e.g., 10, 20 or 40 g/L at least), comprising culturing a mutant microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*) or Gram-negative) having a mutant pantothenate kinase-encoding *coaA* gene, under conditions such that the panto-compound is produced or that production is enhanced, and possibly further recovering the panto-compound. A recombinant microorganism having a mutant *coaA* gene encoding a pantothenate kinase with reduced

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

activity. A recombinant microorganism (e.g., Gram-positive (e.g., belonging to the genus *Bacillus* (e.g., *Bacillus subtilis*), *Corynebacterium*, *Lactobacillus*, *Lactococci* or *Streptomyces*)) that overproduces a panto-compound having a mutation in a *coaA* gene that results in a reduced level of pantothenate kinase activity, resulting in a decrease in the capacity of the microorganism to synthesize coenzyme A. A vector containing a (mutated) *coaA* gene.

## 18. Claim : 94 (partially)

A vector containing regulatory sequences comprising the constitutively active promoter Pveg (SEQ ID NO:41).

## 19. Claim : 94 (partially)

A vector containing regulatory sequences comprising the constitutively active promoter P15 (SEQ ID NO:39).

## 20. Claim : 94 (partially)

A vector containing regulatory sequences comprising the constitutively active promoter P26 (SEQ ID NO:40).

## 21. Claim : 96 (partially)

A vector containing regulatory sequences comprising an artificial RBS according to SEQ ID NO:49.

## 22. Claim : 96 (partially)

A vector containing regulatory sequences comprising an artificial RBS according to SEQ ID NO:50.

## 23. Claim : 96 (partially)

A vector containing regulatory sequences comprising an artificial RBS according to SEQ ID NO:51.

## 24. Claim : 96 (partially)

A vector containing regulatory sequences comprising an artificial RBS according to SEQ ID NO:52.

## 25. Claim : 96 (partially)

A vector containing regulatory sequences comprising an



FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

artificial RBS according to SEQ ID NO:53.

26. Claim : 96 (partially)

A vector containing regulatory sequences comprising an artificial RBS according to SEQ ID NO:54.

27. Claim : 96 (partially)

A vector containing regulatory sequences comprising an artificial RBS according to SEQ ID NO:55.

28. Claim : 96 (partially)

A vector containing regulatory sequences comprising an artificial RBS according to SEQ ID NO:56.

29. Claim : 96 (partially)

A vector containing regulatory sequences comprising an artificial RBS according to SEQ ID NO:57.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US.00/25993

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 0590857	A	06-04-1994	CN 1095105 A,B	16-11-1994
			CN 1225394 A	11-08-1999
			CN 1225388 A	11-08-1999
			CN 1225392 A	11-08-1999
			CN 1225358 A	11-08-1999
			JP 6261772 A	20-09-1994
			US 5518906 A	21-05-1996
-----				
WO 9818954	A	07-05-1998	US 5834259 A	10-11-1998
			EP 0934426 A	11-08-1999
			JP 2001503979 T	27-03-2001
-----				
EP 0224294	A	03-06-1987	AT 56989 T	15-10-1990
			DE 3674548 D	31-10-1990
			DK 533286 A	09-05-1987
			ES 2031818 T	01-01-1993
			FI 864547 A	09-05-1987
			JP 62181789 A	10-08-1987
			PT 83714 A,B	01-12-1986
			US 5024943 A	18-06-1991
-----				
EP 1006192	A	07-06-2000	DE 19855313 A	08-06-2000
			BR 9905776 A	24-04-2001
			CN 1256314 A	14-06-2000
			HU 9904447 A	28-11-2000
			JP 2000228990 A	22-08-2000
			SK 163399 A	11-07-2000
			US 6184007 B	06-02-2001
-----				
EP 1006189	A	07-06-2000	DE 19855312 A	08-06-2000
			BR 9905783 A	24-04-2001
			CN 1256313 A	14-06-2000
			HU 9904448 A	28-11-2000
			JP 2000166580 A	20-06-2000
			SK 164099 A	11-07-2000
			US 6177264 B	23-01-2001
-----				